The Dissertation Committee for Brian Aki Ikkanda Certifies that this is the approved version of the following dissertation:

ASSEMBLY OF COMPLEMENTARY NAPHTHYL UNITS IN NUCLEOTIDOMIMETIC FOLDAMERS

Committee:

Brent L. Iverson, Supervisor

Eric V. Anslyn

C. Grant Willson

Adrian T. Keatinge-Clay

Sean M. Kerwin

ASSEMBLY OF COMPLEMENTARY NAPHTHYL UNITS IN NUCLEOTIDOMIMETIC FOLDAMERS

by

Brian Aki Ikkanda, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2016

Dedication

The heavens declare the glory of God,

And the sky above proclaims his handiwork.

~ Psalm 19:1

This work is dedicated to

my mom, dad, sister, wife and daughter,

all of whom I love dearly.

Acknowledgements

There are too many to thank for the completion of this work but I will try my best to address all who ought to be thanked.

First and foremost, I would like to thank our God and Father of our Lord Jesus Christ, from whom the rational study of science is made possible. He is indeed the Great Creator and Masterful Architect of the universe. I thank Him for saving me and bringing me to UT where I learned much about my identity as a minister of the Gospel but also as a child of God.

I would also like to thank Dr. Brent Iverson who has been an incredible boss, counselor, mentor, and teacher. I owe so much to your guidance not only through my research but also through cultivating my passion for teaching. I hope I can continue the legacy of passionate, dedicated teaching.

Thank you mom, dad and Emi. You have been incredibly supportive since day one and I cannot thank you enough for being there all my life, supporting me, teaching me to be a man of integrity and to be a thinker. I am thankful that you have always been supportive of my passion for chemistry, even getting me that chemistry set in elementary school. But more than that, I thank God for your love and care all throughout my life.

I would like to thank Joyce and Brookie. You have been such a source of strength and encouragement and I am privileged to be so blessed with a wonderful wife and daughter. Through all the sacrifice that you had endured for me to study in graduate school, you have always been supportive and I am forever grateful for that.

To my Gracepoint family who also has been a primary reason for coming to study at UT in the first place. To my spiritual leaders, Pastor Jonathan and Susanna in

particular, for seeing me through my toughest times in graduate school; but also to my previous spiritual leaders Pastor Manny and Sunny, Pastor Will and Esther, and Ray Wong, without whom, I do not know if I would even be out at UT or continuing to fight the good fight.

To the Iverson Lab, both past and present, I must express my gratitude. For the former members Chelsea, Stevan, Amy and the Mike's, thank you for all your guidance and encouragement when I needed help or advice. Cam, it has been a pleasure to bounce ideas off of each other. Liz and Maria, thank you for your 100% help and encouragement, always willing to help get the job done when things got stressful. Chris, your passion and hard work is most impressive and I thank you for not only helping out in the lab and TAing but also for our conversations about life and beyond, I wish you the best in whatever path you decide. A special thank you to Chris and Liz for all the help and support especially in these last few months of finishing up my graduate work, you both are awesome!

To my peers (both the Austin '08/'09-ers and my 2009 peers spread over the world) thank you for your prayers, words of support, and even sitting through practice talks that may not have made any sense to you.

To those younger in the faith that are close to my heart (our homegroup), the class of 2013/2014/2018 in particular but also all those younger in the faith at our church out in Austin who have been my support and encouragement to keep going; for it is your faith that I pressed on as God has entrusted me.

To the King of the ages, immortal, invisible, the only God, be honor and glory forever and ever. Amen.

~ 1 Timothy 1:17

ASSEMBLY OF COMPLEMENTARY NAPHTHYL UNITS IN NUCLEOTIDOMIMETIC FOLDAMERS

Brian Aki Ikkanda, Ph.D.

The University of Texas at Austin, 2016

Supervisor: Brent L. Iverson

The large and complex architectures found in biomolecules not only are an amazing feat of molecular construction, but they also function to carry out all the processes necessary to sustain a living organism. For the organic chemist and biochemist that wish to mimic, improve or modify the structure and function of such natural molecular architectures, an understanding of non-covalent interactions in solution is necessary. In order to probe the particular inter- and intramolecular interactions involved in constructing higher order molecular architectures, the study of foldamers, small synthetic oligomers that adopt well-defined conformations in solution, has emerged.

Our lab has been studying the particular interaction between complementary aromatic units, 1,4,5,8-naphthalenetetracarboxylic acid diimide (NDI) and 1,5-dialkoxynaphthalene (DAN) and their ability to drive the construction of aromatic foldamers in aqueous environments. Utilizing flexible peptide linked units of DAN and NDI, our lab has constructed a variety of unique folding assemblies. From the very first aromatic foldamer that folds into a pleated secondary structure in water to a heteroduplex system assembling oligo-DAN and oligo-NDI units in an intertwined fashion, the use of these complementary aromatic units has proven its versatility in foldamer assembly.

Generally, this dissertation describes studies of the first NDI and DAN deoxyribonucleic acid oligonucleotides. While our previous foldamer assemblies utilized peptide linkers to string together DAN and NDI units, the work described herein draws inspiration from the sugar-phosphate backbone of DNA to assemble aromatic units.

Chapter 2 describes the design and synthesis of 4,4'-dimethoxytrityl protected DAN and NDI phosphoramidites that can subsequently be incorporated into strands of natural DNA oligonucleotides using automated solid phase oligonucleotide synthesis. Chapter 3 investigates the insertion of a 3 base pair region of varying sequences of DAN and NDI into a 12-mer DNA oligonucleotide to explore the effect of DAN and NDI on duplex structure and stability. Chapter 4 describes the stability and structure of a 9-mer DNA oligonucleotide that incorporates two NDI-DAN-NDI triplet sequences either inserted into the interior of the duplex or appended at the terminal positions of the duplex. The NDI-DAN-NDI triplets appended to the terminal positions led to a profound increase in duplex stability (in comparison to internal positions), significantly beyond that seen with analogous sequences of G-C base pairs.

Table of Contents

List of Tables

List of Figures

List of Schemes

List of Equations

CHAPTER 1

Interactions of Aromatic Units for Folding and Assembly in Aqueous Environments

1.1. Molecular Biological Assembly: Complex Architectures from Simple Building Blocks

A brief peek into the world of biological activity at a molecular level will soon have the observer marveling at the intricate array of large complex molecular architectures, each with a particular function and role to sustain life. For example, adenosine triphosphate (ATP) synthase is a multicomponent enzyme of about 500 kDa that acts as a molecular rotor to catalyze the formation of ATP, the energy unit of cells (Figure 1.1a). Deoxyribonucleic acid (DNA), the genetic material of all life, maintains a regular double helical structure over hundreds of millions of base pairs in a single chromosome (Figure 1.2a). Within DNA is all the genetic information to be translated into proteins that serve all the biological function within the cell. And yet, one of the most surprising facts of biology on a molecular level is that the fundamental components of such large and intricate architectures are not all that complex. The molecular components of all proteins can be reduced to varying linear sequences of a mere 20 unique naturally occurring amino acid residues (Figure 1.1b). The building blocks of DNA consist of just 4 nucleobases appended to a sugar phosphate backbone (Figure 1.2b). From strands of one-dimensional information (various sequences of monomer units linked along a oligomer or polymer chain), are produced complex three-dimensional architectures and functional assemblies.

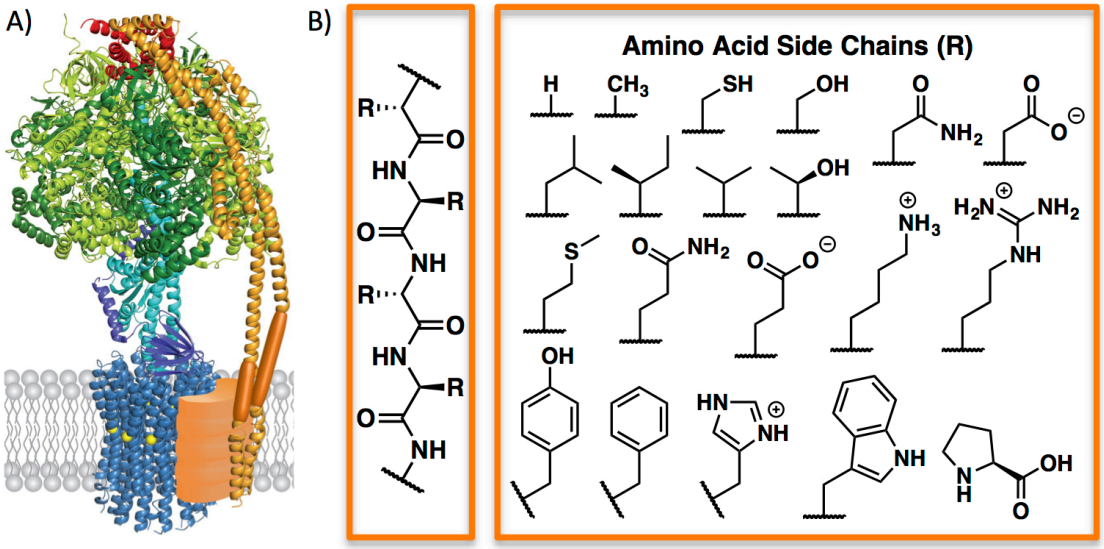


Figure 1.1 (A) A ribbon representation of ATP synthase with the (B) linear peptide chain molecular structure along with the 20 naturally occurring amino acid side chain building blocks.[1]

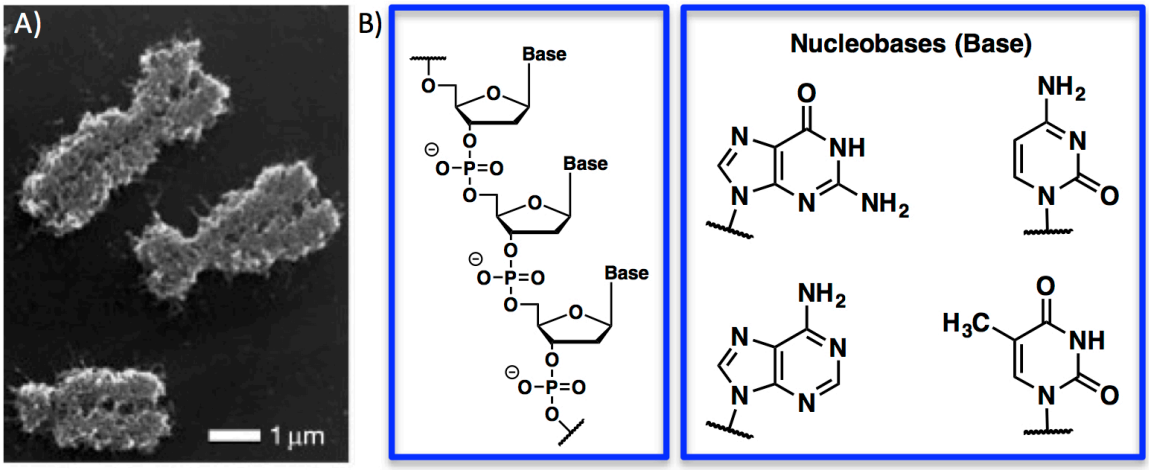


Figure 1.2 (A) A chromosome micrograph along with the (B) linear sugar phosphate backbone and the four nucleobases. Adapted from Methods in Cell Biology, 88, Wanner, G. and Schroeder-Reiter, E. Scanning Electron Microscopy of Chromosomes, 451-474, Copyright 2008, with permission from Elsevier.

A well-established principle in biological chemistry is that the *structure* of large, complex biomolecules is crucial for their *function*. Without well-defined structures, the particular reactions and processes within a living cell will utterly break down. The binding of specific substrates or the opening/closing of channels depend on particular molecular conformations and interactions of structures. The intricate array of competing and interacting non-covalent interactions (hydrogen bonding, ionic interactions, van der Waals, and especially desolvation in water, to name a few) constitute the driving force for the folding and assembly of biochemically useful structures (Figure 1.3). These non-covalent *inter-* and *intra*molecular interactions introduced by different combinations of monomeric units along an oligomer or polymer chain are able to precisely form complex higher order molecular architectures such as the DNA double helix or protein secondary, tertiary, and quaternary structure.

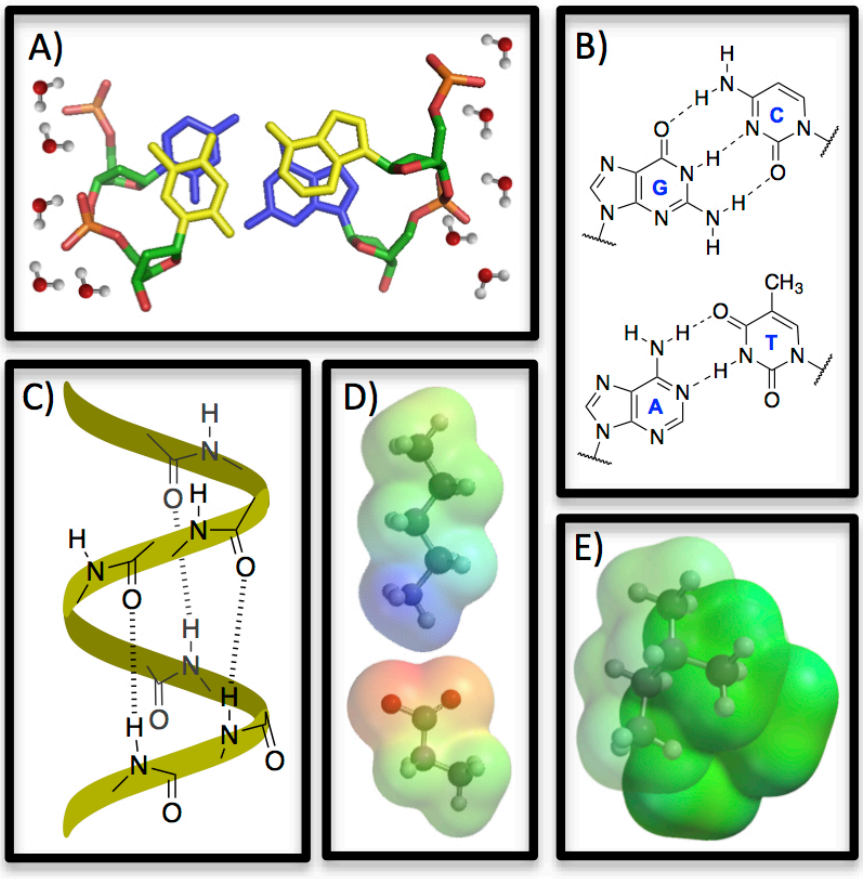


Figure 1.3 Various non-covalent interactions involved in higher order molecular architectures in proteins and DNA. (A) Solvation/Desolvation. (B) Watson-Crick hydrogen bonding patterns. (C) Hydrogen bonding stabilizing α-helical secondary structure. (D) Ionic interactions. (E) van der Waals interactions.

How these non-covalent interactions work together to form such complex and highly ordered architectures is of particular interest for organic chemists and biochemists who desire to improve, mimic, or manipulate natural biological systems. This work focuses attention on non-covalent interactions that dictate folding to drive predictable and designed synthetic assemblies.

1.2. Self-Assembly of Folding Molecules

To try to understand the principles and components of higher order biomolecular architectures, it would seem rational to break them down into their component parts. Therefore, the field of "foldamer"[2] chemistry emerged, which simplified the large biological architectures into smaller oligomers that "fold" upon themselves in solution. More specifically, as defined by Moore and coworkers, a foldamer is "any oligomer that folds into a conformationally ordered state in solution, the structures of which are stabilized by a collection of noncovalent interactions between nonadjacent monomer units. There are two major classes of foldamers: single-stranded foldamers that only fold (peptidomimetics and their abiotic analogues) and multiple-stranded foldamers that both associate and fold (nucleotidomimetics and their abiotic analogues)."[3] Therefore, foldamers provide a two-fold approach to understanding large supramolecular architectures: a top-down approach, taking inspiration from folding molecules in nature and designing oligomers that exploit similar folding principles, and a bottom-up approach, which exploits novel non-covalent interactions (or combinations thereof) to drive folding assemblies. By synthesizing oligomers that are designed to fold in particular conformations in solution, chemists are able to identify the components of higher order self-assembly in solution. These oligomers that fold in solution have given tremendous insight about structure and function of natural and abiotic higher order architectures. Some key examples will be discussed to demonstrate the breadth of study as well as applications that have come about since the inception of foldamer chemistry.

1.2.1. Non-Natural α-, β-, γ-, and δ- Amino Acid Foldamers

Taking inspiration from natural α-amino acid architectures like the α-helix and the β-pleated sheet, Gellman *et al*.[4-6] and Seebach *et al*.[7-10] constructed and characterized

helical and sheet architectures from oligomers of non-natural α-, β-, γ-, and δ- amino acids. Their work focused not on the side chains of amino acids to drive assembly, but rather altering the backbones of oligomers to drive the formation of well-defined secondary structures. A series of foldamers have been designed and synthesized by their respective groups to produce predictable higher order architectures. Providing rigidity through conformationally constrained cyclic amino acids also proved to form regular helical structures.[11] Additionally, well-defined helical architectures could be achieved through even heterogeneous oligomers containing various repeating sequences of α-, β-, and γ-amino acids (Figure 1.4).[12]

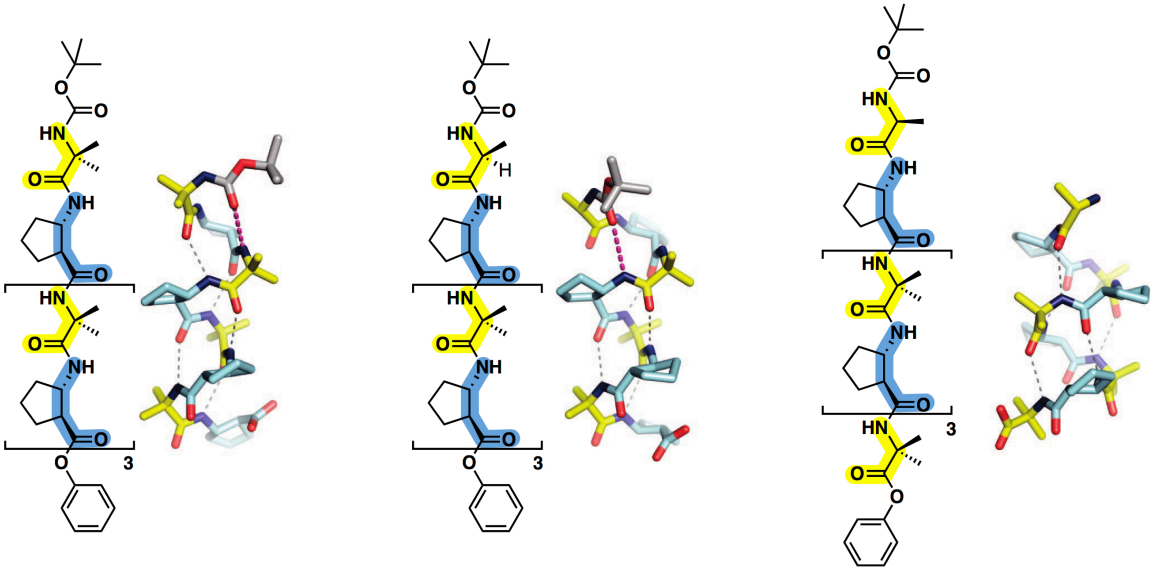


Figure 1.4 Representative α/β-peptides molecular structures and their corresponding crystal structures with highlighted hydrogen bonding patterns. Adapted with permission from Reference 12. Copyright 2008 American Chemical Society.

Recent advances in these modified backbone foldamers have been able to achieve even higher orders of molecular architecture achieving quaternary structure from α/β-peptide helices.[13] Additionally, recent work by Gellman and coworkers has demonstrated

biological applications of these foldamers *in vivo* by successfully mimicking analogous "α-peptides" in blocking protein-protein interactions.[14]

1.2.2. Aromatic Oligoamide Foldamers

Lehn *et al*. focused on hydrogen bonding patterns within amide linked polyheterocyclic oligomers. Through varying sequences of amide-linked azaheterocycles, Lehn *et al*. has established several helical folded conformations.[15-17] Intramolecular hydrogen-bonding of amide hydrogens to the neighboring pyridine nitrogen is designed to stabilize the bent conformation that will form helices in solution in their oligomeric form. Taking this approach, a family of oligomeric analogues have been synthesized to form double helical dimers in solution (Figure 1.5).[18] Since then, aromatic oligoamides have demonstrated their versatility in the construction of hydrogen-bond directed foldamers achieving in recent years several varieties of architectures such as helices, linear tapes, catenanes and rotoxanes, to name a few.[19]

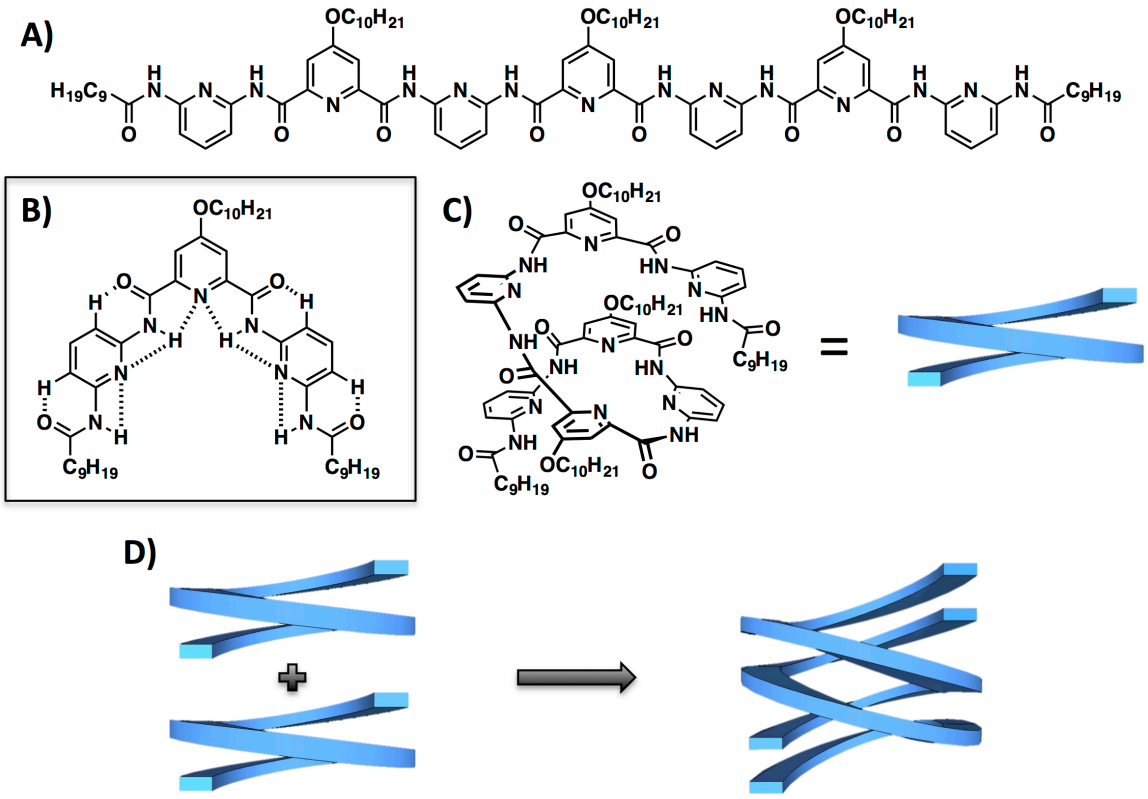


Figure 1.5 (A) Molecular structure of an oligoamide that adopts a helical conformation. (B) Short oligomer highlighting the hydrogen-bonding pattern for the bent conformation. (C) The oligomer adopts a helical conformation in solution and has been shown to (D) Cartoon demonstrating the assembly into a double helix in solution.[18]

Recently, Huc *et al*. synthesized and characterized aromatic oligoamides that fold into multi-stranded artificial β-sheet architectures utilizing a rigid 4,6-dinitro-1,3-phenylenediamine turn unit (Figure 1.6).[20] This particular design differs from natural β-sheet architectures by focusing on the interaction between aromatic units in organic solvents.

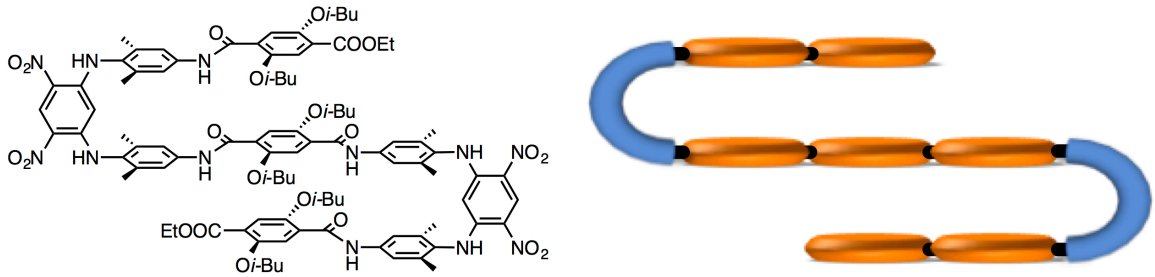


Figure 1.6 Aromatic oligoamides that fold into β-sheet structures (*left*) molecular structure and (*right*) cartoon of folded conformation.

1.2.3. *meta*-Phenylene Ethynylene Foldamers

By taking advantage of solvation/desolvation, research groups have designed oligomers that will fold upon being placed in strongly interacting polar solvents. Moore and co-workers have developed stable helical structures from the solvophobically driven folding of *meta*-phenylene ethynylene oligomers.[21] These helical structures have successfully been able to fold in solution to encapsulate hydrophobic guests within its hydrophobic cavity bearing resemblance to an enzyme holding its substrate in its reactive pocket. These helical architectures have also demonstrated sequence-specific binding of a piperazinium dihydrochloride salt giving greater insight into foldamer design in binding phenomena (Figure 1.7).

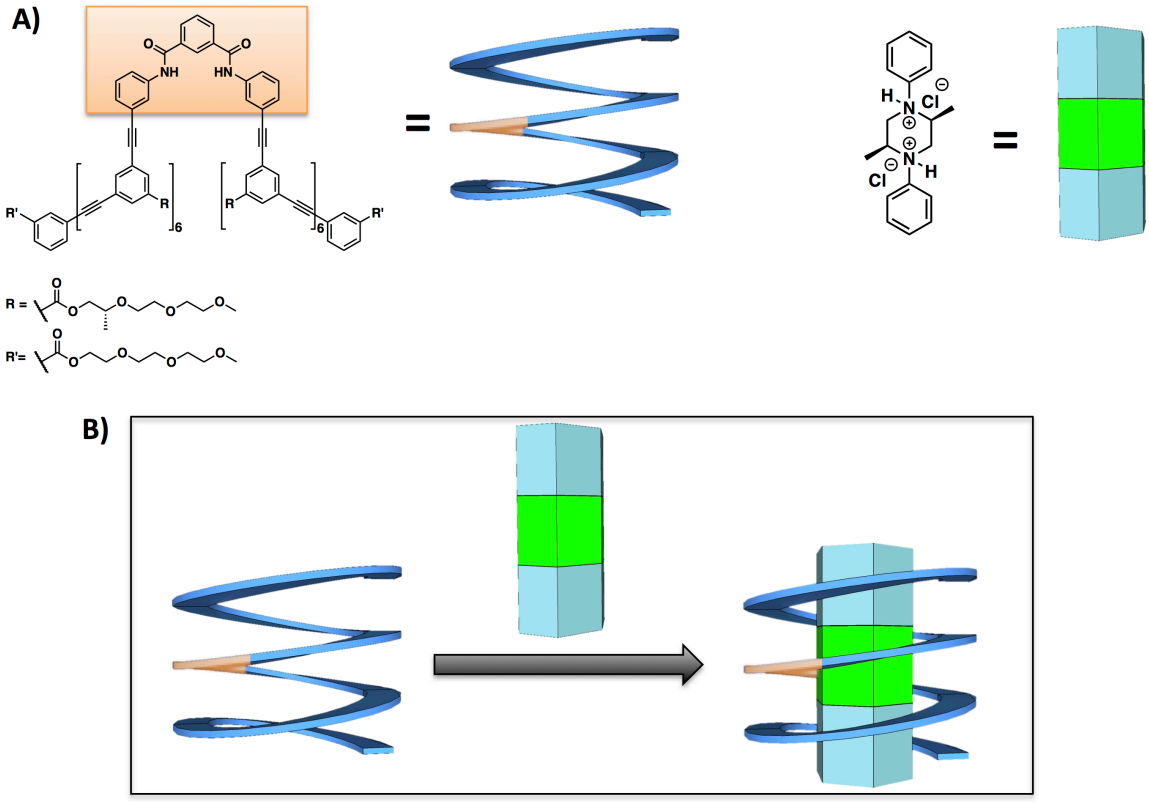


Figure 1.7 (A) Molecular structures of (*left*) solvophobically-driven foldamer and (*right*) piperazinium dihydrochloride ligand with cartoon representations. (B) Scheme of helical foldamer binding its ligand.[22]

The design and synthesis of foldamers that can encapsulate small molecules provides a synthetic scaffold for potential enzyme-like function. These foldamers have been found to be stable in a variety of polar and nonpolar solvents (including DMSO, ethyl acetate, acetonitrile, and methanol for example).

1.2.4. Peptoid Foldamers

Zuckermann and coworkers have had tremendous success generating folded architectures from *N*-substituted glycine oligomers, deemed "peptoids". Void of the chirality at the alpha carbon position, these oligomers surprisingly have been utilized to

assemble into helical structures in solution.[23] More recently, peptoid oligomers of two lengths of alternating cationic/aromatic functional groups and alternating anionic/aromatic functional groups have been demonstrated to form well-defined and robust nanosheet architectures in aqueous environments.[24] These nanosheets have been utilized to provide a platform to develop nacre mimetic materials by being mineralized with thin films of amorphous calcium carbonate, demonstrating versatility in foldamer applications (Figure 1.8).

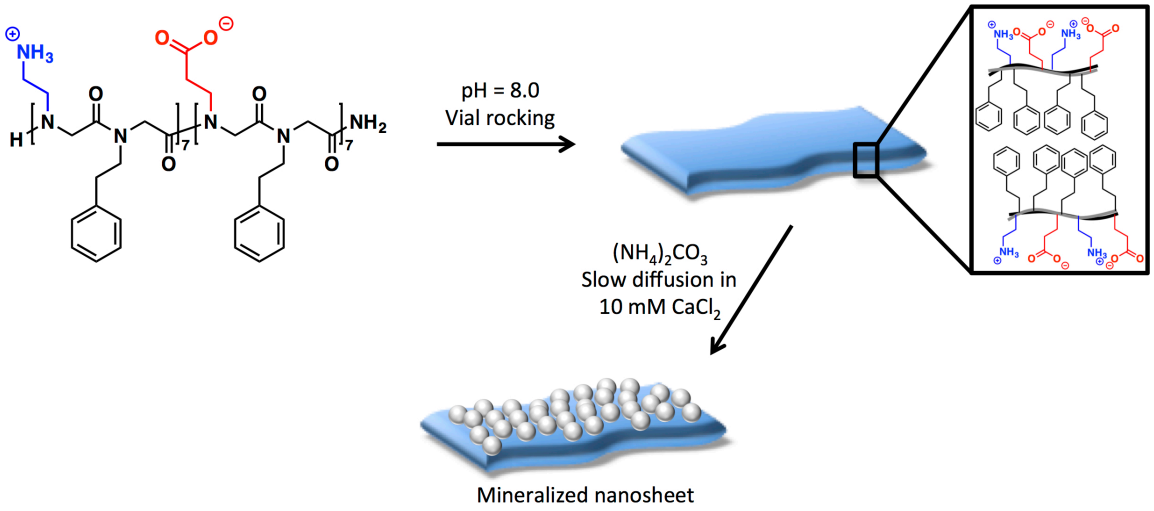


Figure 1.8 Example of sheet construction from amphiphilic peptoids that is subsequently mineralized with calcium carbonate.[24]

The ability of peptoid oligomers to assemble into well-defined sheet architectures provides a platform for functional monomers to be incorporated into the sheet surface or interior. Functionalizable peptoid nanosheets may provide a new scaffold for molecular recognition as particular ligands or protein substrates can potentially be displayed on the surface of the nanosheet.

1.2.5. Nucleotidomimetic Foldamers

Although not frequently cited in terms of foldamer chemistry, nucleic acid related oligomers that have unnatural backbones or unnatural nucleobases that assemble into a double helical heteroduplex conformation in solution fall under the definition of a foldamer. These nucleotidomimetic foldamers either draw their inspiration from the double helical architecture or utilize the well-defined double helical structure of DNA to assemble molecules (often chromophores). Exploration in this arena has led to the development of non-natural DNA backbones that can maintain double helical architecture such as PNA, GNA, and LNA as examples (Figure 1.9).

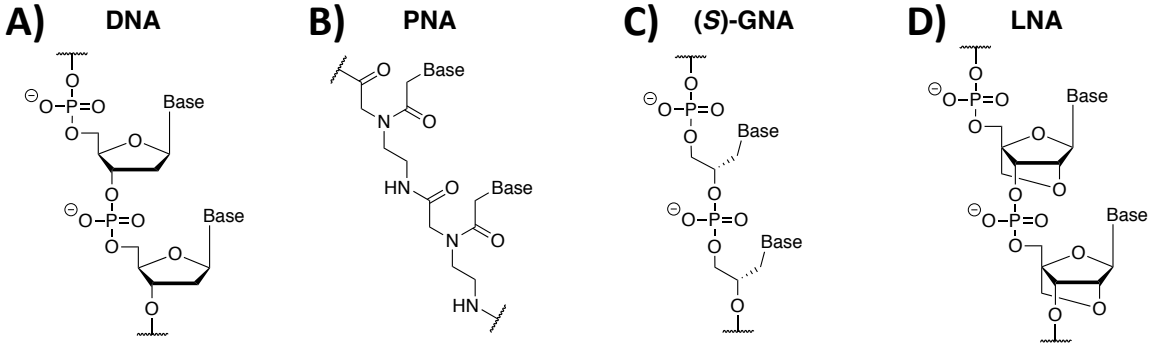


Figure 1.9 General structures of natural (A) deoxyribonucleic acid, DNA, (B) peptide nucleic acid, PNA, (C) glycol nucleic acid, GNA, (D) locked nucleic acid, LNA.

Also, in an effort to expand the genetic code of life from the 4 nucleobases to potentially new non-natural DNA base surrogates, several groups have developed "base pairs" that can actually be enzymatically incorporated into a DNA double helix while conserving the double helical structure (Figure 1.10). The advancement of this area has inspired much of the progress described in this work.

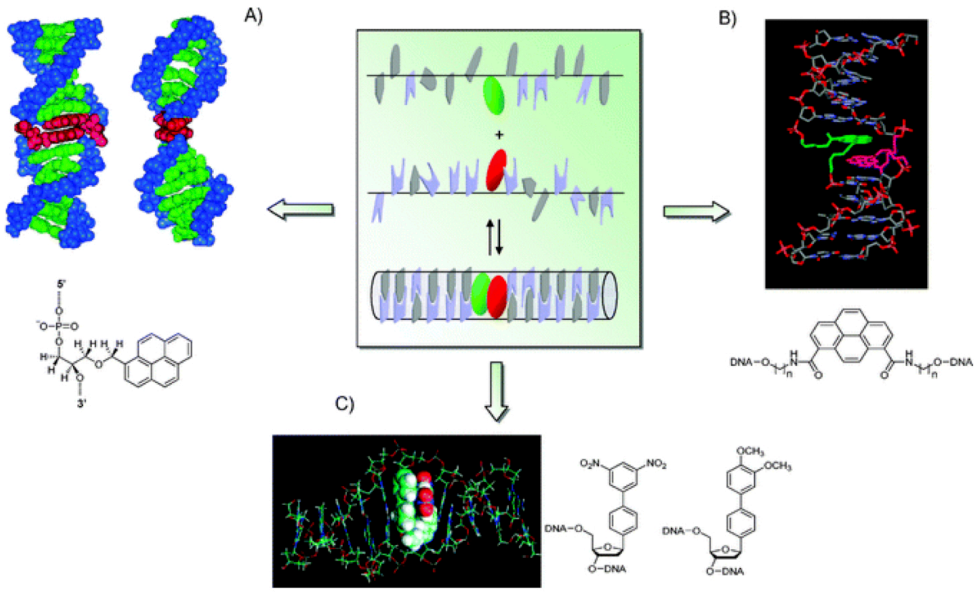


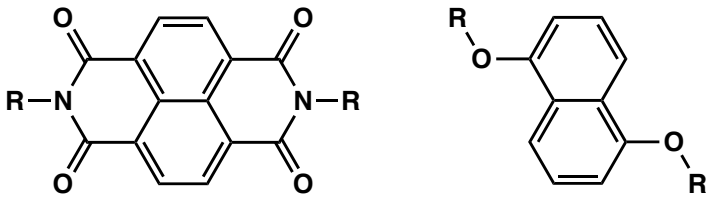
Figure 1.10 Selected examples of non-natural DNA base surrogates incorporated into oligonucleotides. Reproduced from Reference 25 with permission of the Royal Society of Chemistry.[25]

Abiotic nucleotidomimetic heteroduplexes have also been constructed utilizing various other non-covalent interactions to fold in a variety of architectures. Hydrogen bond-driven heteroduplex formation has been demonstrated by Gong and co-workers in the formation of linear oligoamide duplexes[26] and also Krische and co-workers in the construction of duplex forming oligoaminotriazines[27] for example. Combinations of both aromatic interactions and hydrogen bonding were exploited in the formation of hetero duplexes formed by Bisson and co-workers in a zipper-like fashion[28] and also by Lehn and co-workers in a double helical conformation.[18, 29] Additionally, Lehn and co-workers have developed metal-coordinating foldamer heteroduplexes of a helical conformation.[30-

[31] In sum, a variety of non-covalent interactions (and combinations thereof) have been utilized in the construction of foldamers of various architectures and functions.

1.2.6. Complementary Aromatic Unit Foldamers in Water

Our lab has undertaken the study of the specific non-covalent interactions between aromatic units, particularly in the study of 1,5-dialkoxynaphthalene (DAN) and 1,4,5,8-naphthalenetetracarboxylic acid diimide (NDI) units in the construction of foldamers and other assemblies in aqueous environments (Figure 1.11). Based upon early work demonstrating that the relatively electron-rich DAN and relatively electron-deficient NDI associate in aqueous solution, the first aromatic foldamer operating in water (called an "*aedamer*" at the time, *A*romatic *E*lectron *D*onor *A*cceptor oliga*mer*) was designed and synthesized by linking DAN and NDI units in an alternating fashion with flexible peptide linkers.[32] The successful assembly of the well-defined pleated architecture of this aromatic foldamer laid the groundwork for several other foldamer designs as well as efforts to use aromatic interactions to drive aromatic supramolecular assembly in our lab and elsewhere. Since then, the understanding of aromatic interactions have been further refined through computational and experimental work, which has enabled more intentional rational designs of aromatic foldamers and assemblies.



NDI **DAN**

Figure 1.11 Molecular structure of (*left*) 1,4,5,8-naphthalenetetracarboxylic acid diimide (NDI) and (*right*) 1,5-dialkoxynaphthalene (DAN)

In order to describe the design of our particular aromatic foldamers, it is important to first explain the current models describing aromatic interactions as well as the effect of solvation/desolvation on the interaction of aromatic molecules.

1.3. AROMATIC INTERACTIONS

1.3.1. How Do Aromatic Molecules Interact?

The interaction between aromatic units plays an important role in biomolecular and synthetic assemblies. However, what characteristics and features define the interaction between two aromatic molecules? To consider the interaction of aromatic molecules it is important to consider the electrostatic character of aromatic molecules, the geometry of interactions as well as the interactions of aromatic molecules in solution. As discussed below, all three of these factors influence the interaction of aromatic molecules.

1.3.2. "Polar/Pi" Model

A model to predict aromatic stacking geometry was proposed by Hunter and Sanders in the 1990's that focused on the overall polarization of the aromatic π-electron cloud.[33] By treating the electron rich π-cloud of aromatic molecules as separate from the σ-framework, they proposed a model where the partially positively-charged σ-framework around the periphery of the molecule is "sandwiched" between two partially negatively-charged π-electron clouds (Figure 1.12a), revealing a quadrupole moment when visualized from a side-on perspective. The geometry observed for interacting aromatic molecules are therefore predicted by favorable or unfavorable interactions of the "aromatic core" (the π-electron clouds) and the periphery of the aromatic molecules.

The terminology of the "polar/pi" model was first set forth through the work of Cozzi and Siegel[34-38]. Through a series of experiments utilizing a 1,8-diarylnaphthalene

system, they similarly concluded that the interaction of aromatic molecules can be predicted in terms of the polarization of π-electron density. Both this model and the Hunter-Sanders approach offer an explanation for the general tendency for neutral, non-substituted aromatic molecules, such as benzene or naphthalene, to avoid interaction in a face-centered stacked fashion (Figure 1.12b) due to the electrostatic repulsion of the electron rich π-clouds. Instead, they would rather interact with each other in a T-shaped, or "herringbone", geometry (Figure 1.12c) or possibly an offset stacked geometry (Figure 1.12d), which would maximize favorable electrostatic interactions. The partially negatively-charged π-electron clouds are therefore generally predicted to repel one another, disrupting a face-centered interaction for most aromatic units.

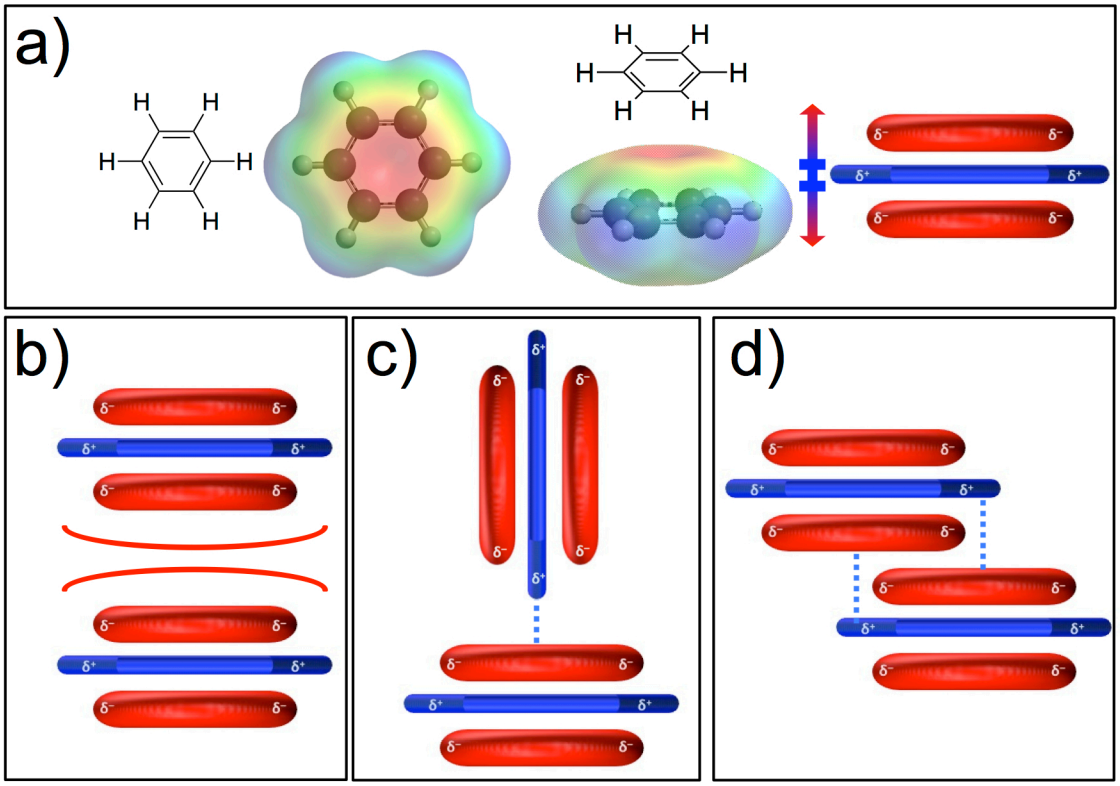


Figure 1.12 (a) Electrostatic potential map of benzene using DFT (B3LYP/6-31G*) as implemented in Spartan (Wavefunction, Inc.), with cartoon depicting quadrupole moment described in the polar/pi model. (b) Repulsion of face centered stacking of benzene. (c) Preferred T-shaped, or "herringbone" interaction of benzene. (d) Somewhat favorable off-centered stacking of benzene.

According to the polar/pi model, the strength of interaction between aromatic units is modulated by the effect of substituents on the electron density of the π-electron cloud. Importantly, this model also predicts that when strongly electron-withdrawing functional groups are placed on the periphery of an aromatic molecule, the π-electron cloud can be polarized such that quadrupole moment is essentially reversed, creating a generally electron-deficient aromatic core and a partially negatively-charged periphery (Figure 1.13a). A reversed polarization of the aromatic core is used to explain why

aromatics with strongly electron-withdrawing substituents prefer to stack in a face-centered geometry when associating with relatively electron rich aromatic molecules. For example, in the solid state benzene and hexafluorobenzene adopt an alternating face-centered stacked geometry (Figure 1.13b).[39-40] This model also predicts that electron-donating substituents will have an opposite effect, thereby increasing the electron repulsion of the π-electron clouds of aromatic compounds in a face-centered stacked geometry.

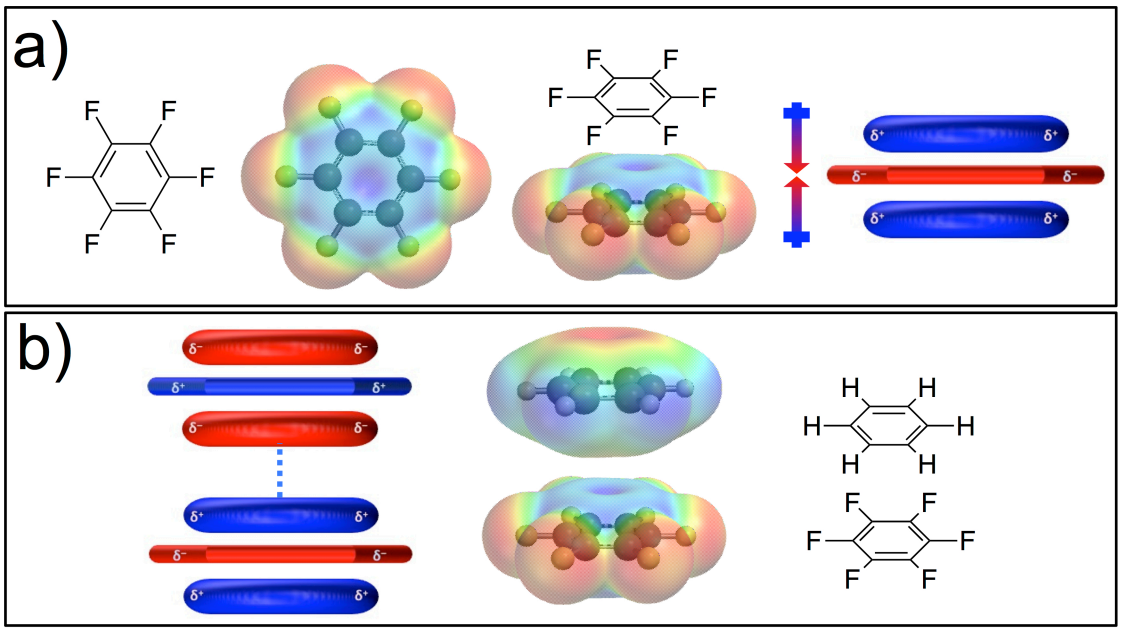


Figure 1.13 (a) Electrostatic potential map of hexafluorobenzene using DFT (B3LYP/6-31G*) as implemented in Spartan (Wavefunction, Inc.), with cartoon depicting its reversed polarization. (b) Favorable face-centered stacking of benzene and hexafluorobenzene.

The polar/pi model therefore rationalizes the strength and geometry of interaction of substituted aromatic molecules through the electrostatic potential maps of π-electron

densities. In particular, the model assumes that electron rich aromatic rings and electron poor aromatic rings are complementary, leading to a favorable stacking interaction.

1.3.3. "Local, Direct Interaction" Model

An alternative view of the interaction between substituted aromatic molecules was suggested through the early experimental work of Rashkin and Waters.[41] Through a model system of *meta*- and *para*-substituted *N*-benzyl-2-(2-fluorophenyl)-pyridinium bromides, they could determine the strength of the stacked aromatic interaction by measuring the rotation barrier about the biaryl bond. Through these experiments, they noted that the *orientation* of stacked substituted aromatic molecules plays a role in the magnitude of observed interactions between aromatic units. It was suggested that direct, through-space electrostatic interactions of polarized substituents along the periphery of the aromatic molecules provide the dominant stabilizing effect.

Using calculated interaction energies between substituted benzene dimers, Wheeler and Houk investigated the idea that direct complementary interactions between the substituents of aromatic units provide a driving force for stacking interactions.[42] For example, their computational studies calculated energies for the interaction between a mono-substituted benzene (C_6H_5-X) and unsubstituted benzene (C_6H_6), which were then compared to the energies calculated for the corresponding H-X analogue with benzene. These studies indicated that no significant additional stabilizing energy was afforded by the aromatic core of the substituted benzene relative to the H-X analogue even when the substituents (X) demonstrated similar trends in interaction energies. This led to the conclusion that "substituent effects in the sandwich configuration of the benzene dimer do not involve the π-system of the substituted benzene".[42]

This observation was further investigated with experimental work by Houk *et al*.[43] and computational work by Wheeler[44-45] to develop a refined model to predict the interaction of substituted aromatic units. Through several computational studies, it was demonstrated that overemphasis on the polarization of the π-electron density may in fact contradict the computed interaction energies whereas focus on the local, direct interactions of substituents seem to provide more comprehensive accurate predictions of computed interaction energies. For example, a strong correlation between the interaction energies of a sandwich dimer of benzene with various monosubstituted benzenes (C_6H_5-X) and the analogous interaction of H-X and propene demonstrates that the interactions of monosubstituted benzene sandwich dimers are dominated by the local interaction of that substituent and the closest vertex of the other aromatic ring (Figure 1.14a).

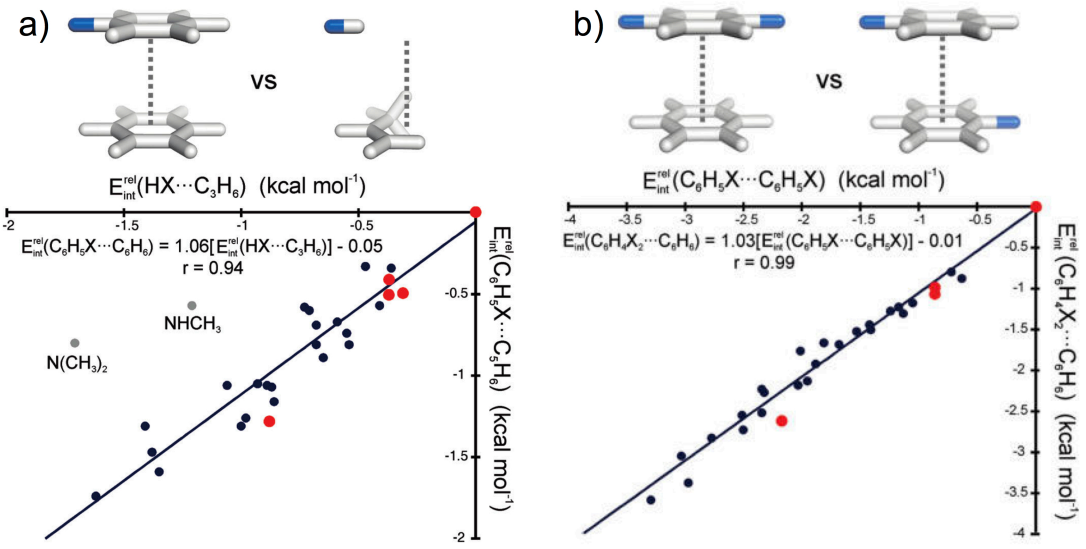


Figure 1.14 (a) Strongly correlated interaction energies (kcal mol^{-1}, relative to X=H) for C_6H_5-X ••• C_6H_6 sandwich dimers versus H-X ••• C_3H_6. (b) Strongly correlated interaction energies (kcal mol^{-1}, relative to X=H) for p-C_6H_4-X_2 ••• C_6H_6 sandwich dimers versus C_6H_5-X sandwich homodimers. Adapted with permission from Reference 44. Copyright 2011 American Chemical Society.

Furthermore, by solely looking at electrostatic potentials of a para-disubstituted ring sandwich dimer with benzene versus the monosubstituted homo dimer (Figure 1.14b) would suggest that the magnitude of interaction energies would be more pronounced by the former case; however, calculated interaction energies reveal strong correlations between the two scenarios. This can readily be explained by the local, direct interactions of the substituents as they would be equal in both cases, rather than the polarization of the π-electron density, which may predict unequal interactions of these two situations. This model, deemed the "local, direct interaction" model,[44] proposes that the local, direct through-space electrostatic attraction/repulsion between highly polarized substituents along the periphery of aromatic units is the dominant factor determining a

stacking geometry. Furthermore, the work of Sherill *et al*.,[46-48] Snyder *et al*,[49] Lee *et al*.,[50] and Grimme[51] all suggest the importance of the interactions of the substituents around the periphery of aromatic molecules.

1.3.4. Interactions of DAN and NDI

Our work in the field of aromatic interactions is based on the use of substituted naphthyl aromatic units (DAN and NDI). The electrostatic potential maps of these two molecules reveal a polarization of the aromatic core derived from the electron-donating alkoxy groups and electron-withdrawing imide carbonyl groups, respectively (Figure 1.15). Therefore, an alternating, face-centered stacked geometry is predicted for DAN and NDI in the solid state (Figure 1.16a).[52] Note that face-centered stacking is predicted by either the "polar/pi" or "local, direct interaction" models.

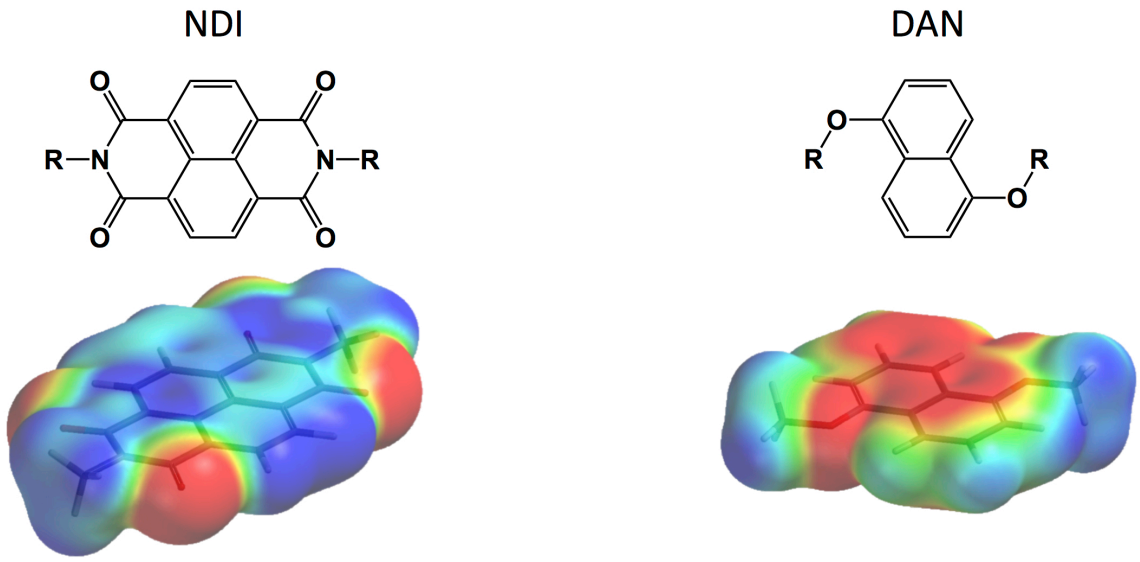


Figure 1.15 Molecular structures and electrostatic potential surfaces calculated for (*left*) NDI and (*right*) DAN using DFT (B3LYP/6-31G*) as implemented in Spartan (Wavefunction, Inc.).

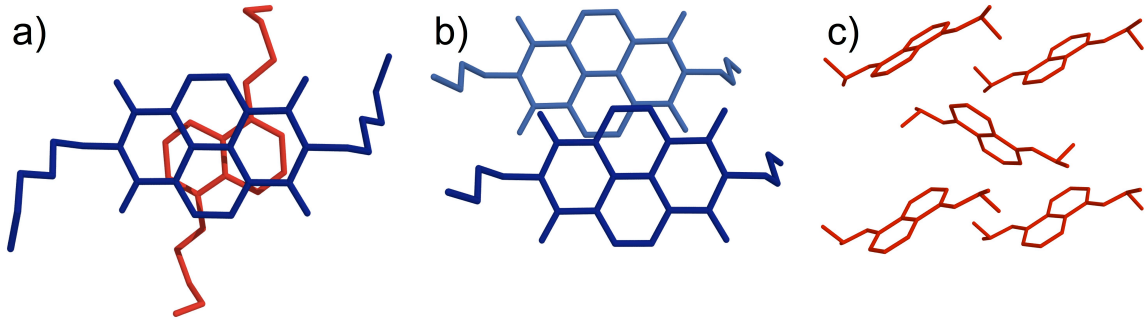


Figure 1.16 X-ray crystal structures of (a) DAN-NDI face-centered stacked monomers, (b) NDI-NDI offset stacked monomers, and (c) DAN herringbone geometry in the solid state.[53-54]

Additionally, it has been observed that the π-orbital mixing introduced by this face-centered geometry results in a charge-transfer absorbance and thus a characteristic red-purple color once the colorless solutions of DAN and NDI are mixed. For situations in which complexes demonstrate charge-transfer absorbance, an alternative aromatic

donor-acceptor terminology has been historically used. The electron-rich aromatics such as DAN are referred to as a "donor" while electron-deficient aromatics such as NDI are referred to as an "acceptor" because the charge-transfer absorbance occurs when an electron from the HOMO of the donor is excited to the LUMO of the adjacent acceptor.

Investigation of the solid state interactions of monomeric DAN and NDI units revealed that a sample of pure NDI crystallizes in an off-set stacking mode (Figure 1.16b). Closer examination of NDI-NDI self-stacked crystal structure reveals that the carbonyl oxygen atom of one NDI is in close proximity to the carbonyl carbon of the adjacent NDI unit, suggesting a qualitative explanation for the solid-state geometry that is best predicted by the "local, direct interaction" model. Note that a sample of pure DAN crystallizes in a herringbone arrangement in the solid state (Figure 1.16c), which is consistent with the predicted favorable orientation of electron rich aromatic molecules of the polar/pi model.

1.3.5. Aromatic Interactions in Strongly Interacting Solvents

Thus far, aromatic interactions have been discussed without the consideration of solvent. However, solvation/desolvation have considerable impact on the association of aromatic molecules, especially in aqueous environments. In general, the electrostatic driving forces described thus far are thought to be important, even dominant, in low polarity, weakly interacting solvents. However, in strongly interacting polar solvents, solvophobic effects dominate. More specifically, solvophobic effects describe the tendency for non-polar solutes to cluster within polar solvents. This effect is often cited to be entropically driven, referring to the high entropic cost of solvating a lone non-polar solute by creating an ordered "solvent cage" around the surface of the solute.

In strongly interacting polar solvents, especially water, aromatic molecules will tend to prefer geometries of maximal surface overlap, thereby reducing the amount of surface area in contact with solvent. The important implication is that desolvation considerations alone favor face-centered stacking of aromatic units. Balanced against this preference for a face-centered stacking geometry can be electrostatic interactions favoring other geometries, such as off-set stacking or even a herringbone type of geometry. The result is an interesting interplay of forces in which the electrostatic interactions between associated molecules influence the solvophobic driving force.

In order to probe the influence of solvophobic interactions on aromatic association, monomeric DAN-DAN, NDI-NDI and DAN-NDI interactions were analyzed in solvents of varying polarity (Figure 1.17).[55]

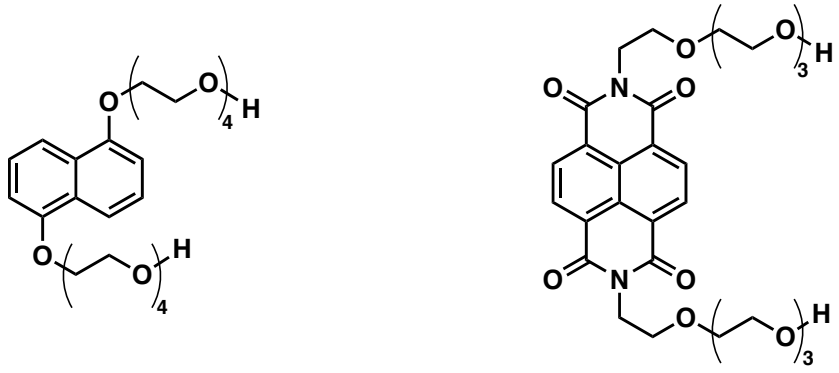


Figure 1.17 Monomer units of (*left*) DAN and (*right*) NDI used to determine association constants in various solvents.

The association constants for the three different combinations of aromatic units were calculated using ^{1}H NMR chemical shift data (Table 1.1). The results demonstrated a significant solvent dependence, with associations being by far the strongest in water.

	Solvent	DAN-DAN K_a (M^{-1})	NDI-NDI K_a (M^{-1})	NDI-DAN K_a (M^{-1})
1	$CDCl_3$	*negligible*	*negligible*	2
2	acetone-d_6	1	1	8
3	DMSO-d_6	1	2	3
4	CD_3CN	1	3	11
5	CD_3OD	1	8	30
6	3:1 CD_3OD/D_2O	1	15	63
7	1:1 CD_3OD/D_2O	2	28	254
8	1:3 CD_3OD/D_2O	10	101	952
9	D_2O	20	245	2045

Table 1.1 Association constants of monomeric DAN and NDI in their self-association or complementary association in various solvents. Adapted with permission from Reference 55. Copyright 2001 American Chemical Society.

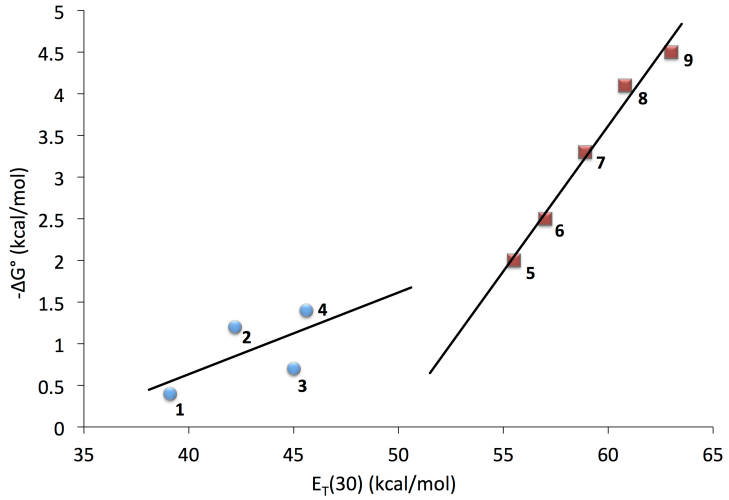


Figure 1.18 The polarity parameter $E_T(30)$[56] plotted against calculated free energy change values for the 1:1 DAN:NDI complexes measured in Table 1.1. Data labels 1-9 represent the solvent conditions from Table 1.1. Adapted with permission from Reference 55. Copyright 2001 American Chemical Society.

When association constants are plotted against the polarity parameter $E_T(30)$[56], a strong correlation with solvent polarity is obvious (Figure 1.18). Notably, the line fit for polar-protic solvents has a somewhat steeper slope than the other organic solvents,

suggesting that hydrogen-bonding is particularly important for aromatic associations. This therefore suggests that the dominant energetic driving force for the association of relatively non-polar aromatic molecules is solvation/desolvation, i.e. the hydrophobic effect. However, this is not the entire story.

Importantly, in water, the DAN-NDI association constant was the strongest by an order of magnitude. The second strongest interaction was seen with the NDI-NDI interactions, while the DAN-DAN complex showed the weakest interaction in every solvent. Solid-state structural preferences, derived largely from interactions of highly polarized substituents on the periphery of the aromatic rings (local, direct interaction model), are predictive of overall association constants in water when the amount of solvent-exposed surface area is considered. The DAN-NDI interaction, known to prefer alternating face-centered stacking in the solid state, shows the highest association constant, presumably because maximum hydrophobic surface area is buried due to the face-centered stacked geometry. The NDI-NDI complex, known to prefer an off-set stacking geometry in the solid state would lead to an intermediate level of hydrophobic surface area exposed, exhibits an intermediate association. The DAN-DAN complex, known to prefer a herringbone arrangement in the solid state that would lead to the most hydrophobic surface area left exposed to solvent, displayed the lowest association constants.

In sum, with regard to recent advances in aromatic interactions from both computational and experimental studies as well as our own contribution to the field, an emerging picture of aromatic stacking interactions is one dominated by the geometrical association that allows for the most favorable interactions of highly polarized substituents along with solvation/desolvation considerations.

You will notice that "π-stacking" and/or "π-π interactions" were not mentioned when describing our work in this field despite the fact that stacked aromatics are being discussed. As described in the recent article by Martinez and Iverson[57], these terms do not have meaning when discussing energetically important interactions despite their common use (or misuse in our opinion) in the literature. When it comes to aromatic units, stacking is a geometrical description, not a particular interaction.

While dispersion does contribute to the stability of stacked aromatic molecules (and is even the dominant attractive contribution to the interaction energy for the benzene dimer in certain configurations as calculated by energy decomposition analysis)[58] it has also been calculated that the favorable dispersion forces seen in the stacked benzene dimer are also equally important in the stacked cyclohexane dimer, suggesting that aromaticity in monomeric units is not a requirement for the stacked geometry. The association of aromatic units of the size involved with most biological and supramolecular systems (one to four fused aromatic rings) is driven primarily by electrostatic attraction combined with desolvation, not any special interactions associated with the presence of aromatic π-electrons on the face of aromatic units.

1.4. Survey Of Iverson Lab Foldamers

1.4.1. Aromatic Foldamer Assembly

In 1995, Lokey and Iverson described the first aromatic foldamer in an aqueous environment (Figure 1.19).[32] The relatively strong association of DAN and NDI in aqueous solution gave inspiration for an oligomer that would fold driven by their stacking interaction. In an alternating fashion, DAN and NDI units were strung together with flexible peptide linkers. Glutamic acid residues were chosen to link the aromatic units

together to impart water solubility and also to prevent intermolecular aggregation. The molecule was designed to adopt a pleated secondary structure in aqueous solution driven by the face-centered stacking of the alternating DAN and NDI units.

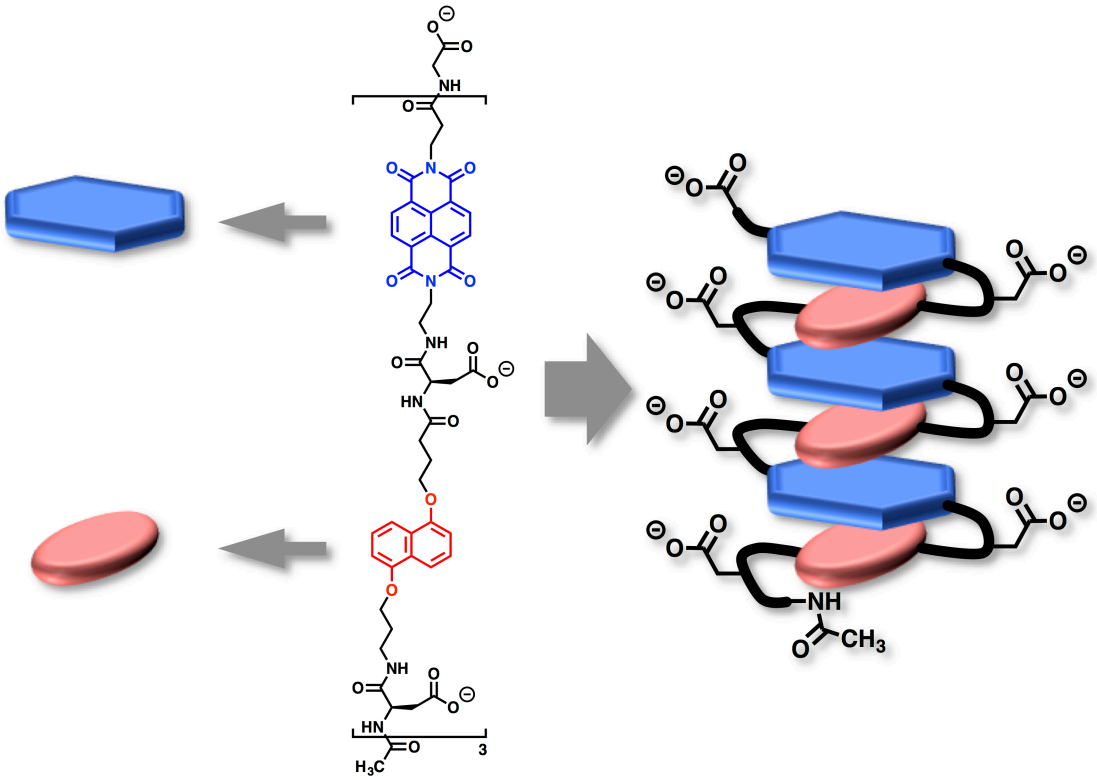


Figure 1.19 Molecular structure and cartoon diagram of folded conformation of Lokey aromatic foldamer.

Comparison of ^{1}H NMR chemical shifts and especially hypochromism in the UV-Vis absorbance spectra (which is characteristic of face-centered aromatic stacking analogously seen in DNA duplex formation) revealed that the DAN-NDI foldamer was indeed folded into the predicted pleated structure, placing the aromatic moieties in a face-centered stacked conformation. The aromatic stacked geometry was further affirmed by

the resulting plum color characteristic of the formation of charge-transfer (CT) absorption bands when DAN and NDI are stacked in a face-centered fashion.

Being peptide-based, this new class of foldamers could be easily synthesized using standard solid phase peptide synthesis (SPPS), which was a favorable feature for exploring the effects of the nature of the linkers on structure and foldamer properties. Additionally, with the use of several different spectroscopic handles (UV-Vis spectroscopy, ^{1}H NMR chemical shifts, and 2-dimensional COSY data) the folding conformation of these aromatic foldamers could be easily confirmed.

The successful formation of the organized architecture gave credence to the pleated secondary structure design generated by aromatic units within an aqueous environment. This first-generation aromatic foldamer has resemblance to the DNA double helix, but is so entirely distinct that it is reasonable to claim an entirely new type of higher order architecture for these folded systems, the first such synthetic system operating in water.

1.4.2. Amphiphilic Aromatic Foldamer

The predictable higher-order architecture and facile synthetic route of our first aromatic foldamer opened the door for deeper investigations of aromatic foldamers, their properties, and the factors that contributed to their ordered structure. Because the linkers that strung together the DAN and NDI units were peptides, the character of the aromatic foldamer could be changed depending on which amino acid residues were chosen. The length of the linker and order of aromatic units could also be changed, leading to aromatic foldamers that were not based on alternating DAN-NDI sequences, yet nevertheless folded in unique topologies that assumed an alternating DAN-NDI stacked arrangement (Figure 1.20).[59]

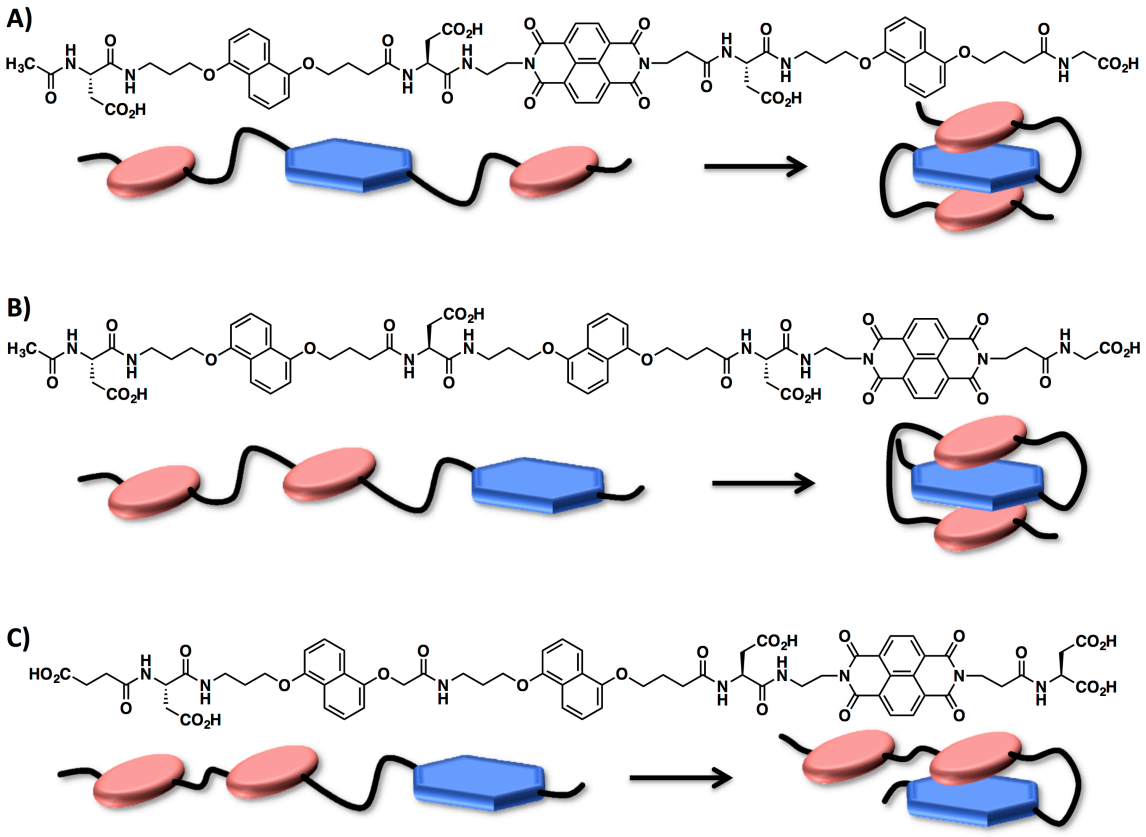


Figure 1.20 Molecular structure and cartoon representation of NDI and DAN trimer foldamers designed to adopt alternative folding patterns: (A) pleated folding, (B) intercalative folding, and (C) trimer stack prevented.[59]

In order to probe the effect of linker character on aromatic foldamer assembly in water, a series of amphiphilic DAN-NDI foldamers were designed and synthesized. When folded, hydrophilic aspartic acid residues were predicted to align down one side of the stacked aromatic column, while a series of selected hydrophobic residues, such as leucine (Figure 1.21), would align down the other side.[60] The structure was designed to be conceptually similar to the leucine zipper motif found in proteins.[61] Spectroscopic analysis of the leucine derivative amphiphilic aromatic foldamer proved to have similar hypochromism and charge-transfer absorbance in aqueous buffer as our previous

aromatic foldamer. However, the unresolved ^{1}H NMR spectra of the molecule in D_2O gave evidence for extensive aggregation, which was further confirmed with dynamic light scattering experiments.

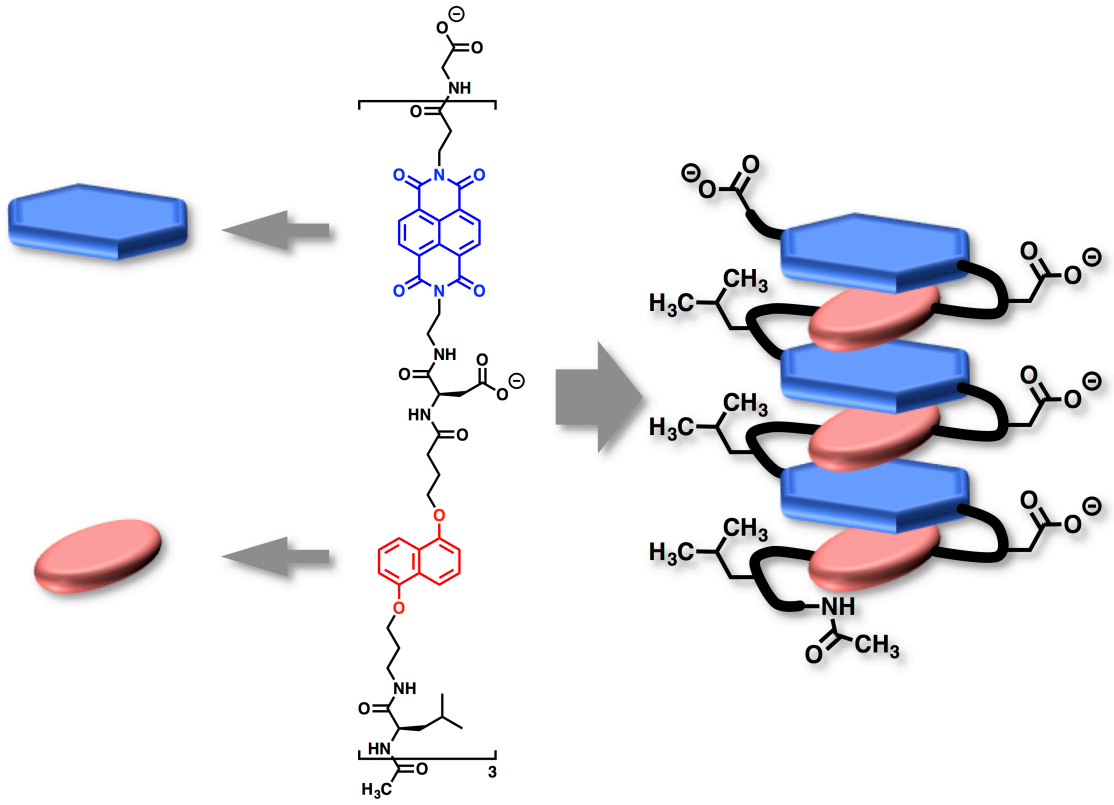


Figure 1.21 Molecular structure and cartoon depicting the folded state of the amphiphilic aromatic foldamer.

Surprisingly, when the aqueous solution of this amphiphilic aromatic foldamer was heated above 80°C, the dark red-purple color changed to a pale-yellow and the viscosity of the solution notably increased. Further spectroscopic investigation of the resulting hydrogel demonstrated the loss of the charge-transfer band, an increased UV-Vis baseline, and a higher degree of aggregation. The hydrogel apparently persisted indefinitely, or at least not returning back to the original state after even months at room

temperature, demonstrating that an irreversible conformational change had occurred upon heating. Importantly, circular dichroism (CD) spectroscopy determined that the aggregated state of the hydrogel was a highly ordered structure. This intriguing behavior not only demonstrated that changing the linkers connecting these aromatic units can have a profound effect on its physical properties, but it also bore striking similarities to the process and properties of natural amyloid formation.[62]

After extensive spectroscopic investigation, a detailed model was proposed for the hydrogel formation of the amphiphilic aromatic foldamer. The work of Parquette *et al*.[63-65], Matile *et al*.[66-67], Govindaraju *et al*.[68-69], and Ghosh *et al*.[70] in NDI-driven supramolecular assemblies, have provided additional insight in the formation of these hydrogel aggregates. Characteristic UV and CD spectroscopic traces indicative of NDI off-set, twisted self-assembly was identified in the aggregated amphiphilic aromatic foldamer sample. Further, transmission electron microscopy (TEM) and atomic force microscopy (AFM) images revealed fibrils of uniform width and regular helicities (Figure 1.22b). Therefore, it was proposed that the formation of this newly aggregated structure was due to a conformational switch of the DAN-NDI *intra*molecular stacked foldamer to a NDI-NDI *inter*molecular off-set stacked fibril aggregate (Figure 1.22a).[71] The NDI-NDI interactions form the basis for an amphiphilic tape-like structure, and two of these come together to occlude the hydrophobic faces of each. The result is a bilayer tape-like assembly, with only hydrophilic surfaces exposed on either side. The amount of hydrophobic surface area buried by this bilayer assembly is extensive, revealing the likely thermodynamic driving force for its irreversible formation. Being the first synthetic analog with behavior analogous to the natural amyloid fibrils observed in many diseases

including Alzheimer's disease, the amphiphilic aromatic foldamer has shed considerable light on the energetics and even mechanism of amyloid fibril assembly.

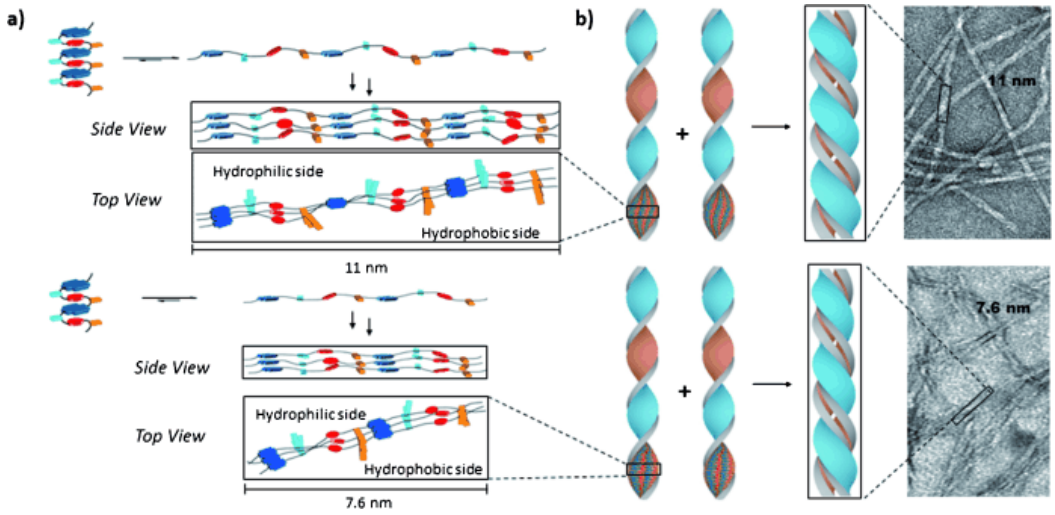


Figure 1.22 Proposed model depicting the conformational shift to form one-dimensional fibrils. (a) Cartoon representation of the conformational shift. (b) The resulting aggregated helical structure and corresponding TEM images. Reproduced from Reference 71 with permission from John Wiley and Sons.

The amphiphilic foldamer studies, along with additional support from other NDI-driven supramolecular structures have confirmed that the NDI-NDI off-set stacked interaction should also be considered as a significant interaction, capable of forming the basis of higher order architectures in aqueous solution.

1.4.3. Heteroduplex Assembly

The aromatic interactions that drove the formation of pleated architectures in our early foldamer designs were then exploited to drive two oligomers (oligo-NDI and oligo-DAN) into a complex heteroduplex structure in aqueous solution (Figure 1.23).[72] Intermolecular interactions of oligomers are of considerable importance for DNA double helix formation as well as protein quaternary structure assembly. These interactions have

been the inspiration for the design of duplexes in the realm of nucleotidomimetic foldamers. The designed architecture of this novel foldamer would require an intertwined conformation (as opposed to a zipper-like conformation found in DNA). These aromatic units were linked together using aspartic acid residues, which function to impart water solubility and reduce the aggregation of like-charged chains, similar to the construction of our first aromatic foldamer. Similar interactions of complementary aromatic units have also been demonstrated to drive heteroduplex formation in organic media by Li and co-workers in both zipper[73] and intertwined[74] conformations.

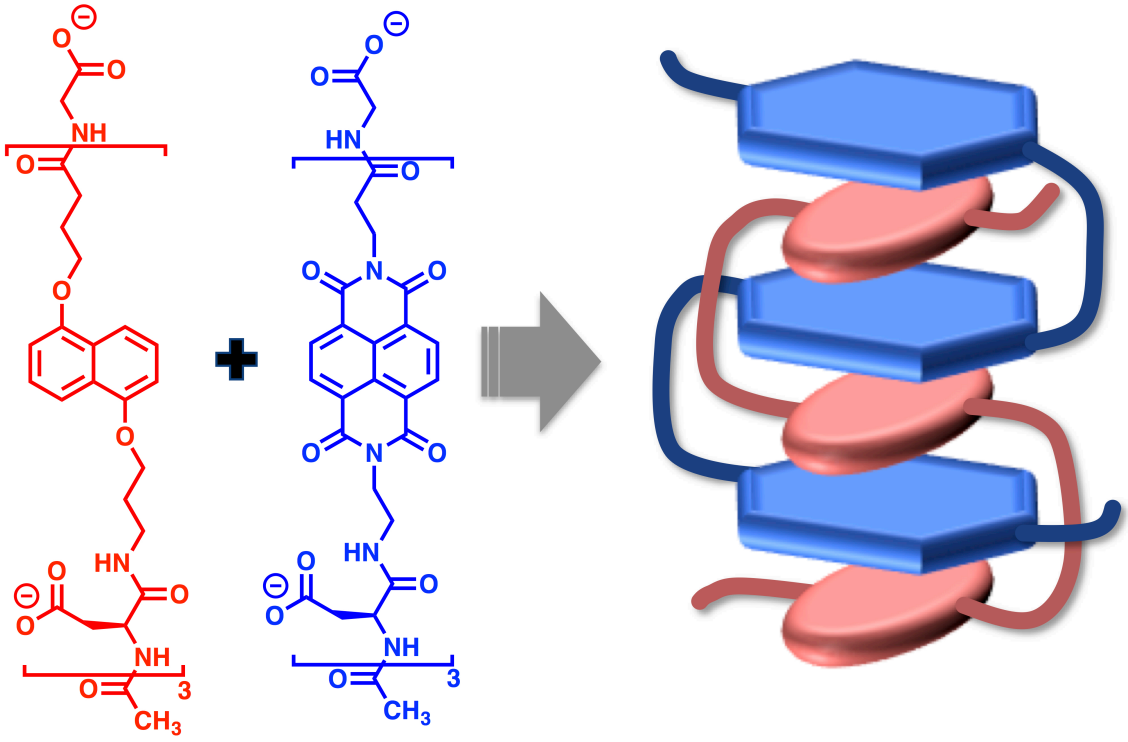


Figure 1.23 Molecular structure and cartoon depicting the formation of the intertwined heteroduplex.

Heteroduplex formation was supported by a series of NMR, polyacrylamide gel electrophoresis (PAGE) and isothermal titration calorimetry (ITC) experiments (Figure

1.24). The expected face-centered NDI-DAN interaction was confirmed through ^{1}H NMR peak shift experiments by comparison with the association of the monomeric DAN and NDI units. A 1:1 duplex formation was confirmed by the PAGE gel experiments in analogy to DNA double helix formation, as a fast-moving oligo-DAN-oligo-NDI duplex was observed as the only band when a 1:1 mixture of the oligo-DAN and oligo-NDI units were mixed. Additionally, a surprising temperature independence in the association constant for the NDI-dimer with its complementary DAN-dimer was observed (Figure 1.24d), indicating the presence of enthalpy-entropy compensation.

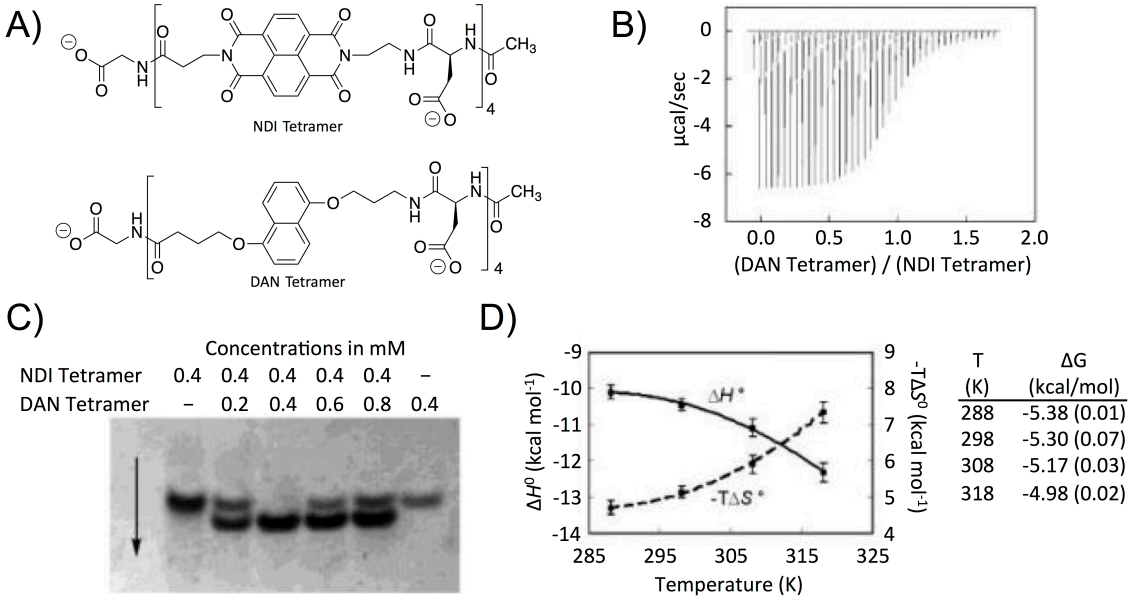


Figure 1.24 (A) Structure of tetrameric oligo-NDI (B) Representative ITC in the association of tetrameric oligo-DAN and oligo-NDI. (C) Gel-shift assay to demonstrate duplex formation. (D) Thermodynamic data of DAN-dimer and NDI-dimer association. Adapted with permission from Reference 72. Copyright 2002 American Chemical Society.

The successful assembly of our complementary aromatic unit-driven heteroduplex opens the door for utilizing complementary aromatic interactions to drive intermolecular assembly in novel nucleotidomimetic foldamer designs.

1.5. OUTLOOK

The electrostatic complementarity of DAN and NDI units have proven to be of unprecedented versatility for the assembly of various aromatic foldamers and other assemblies in aqueous environments. Folding and assembly based on the DAN-NDI interaction has ranged from a pleated secondary structure to higher order architectures of one-dimensional fibrils, and ultimately the assembly of heteroduplexes. Not mentioned here is our extensive work with DNA-binding oligo-NDI molecules.[75-77] Because the stacking of aromatic units is driven in part by polar solvents, one can predict a very bright future for the further development of new generations of aromatic-based foldamers and more complex assemblies in water. Described herein is the advancement of assembly of DAN and NDI in the context of nucleotidomimetic foldamer architectures. Chapter 2 describes the synthesis of novel DAN and NDI phosphoramidites that can subsequently be incorporated into strands of DNA via automated solid phase oligonucleotide synthesis. Chapter 3 describes the insertion of DAN and NDI in various sequences within a 12-mer oligonucleotide duplex and the resulting effect on duplex stability and structure. Chapter 4 describes the position dependence of an NDI-DAN-NDI triplet sequence on DNA duplex thermal stability when two such triplet sequences are placed at an internal position or at terminal positions of a DNA duplex.

This chapter adapted with permission from Ikkanda, B.A. and Iverson, B.L. (2016). Exploiting the Interactions of Aromatic Units for Folding and Assembly in Aqueous Environments. Chem. Commun.

CHAPTER 2

Synthesis of NDI and DAN Phosphoramidite Monomers

2.1. CHAPTER SUMMARY

2.1.1. Introduction

The Iverson lab has demonstrated tremendous success with folded assemblies of oligomeric NDI and DAN units connected with flexible peptide linkers, which adopt a unique pleated or intertwined secondary structure. Because these foldamers exhibit many structural features similar to deoxyribonucleic acid (DNA), we were curious to explore our DAN and NDI interactions within the context of oligonucleotides. In order to create an oligomeric system to organize DAN and NDI within the context of nucleotidomimetic foldamers, automated solid phase oligonucleotide synthesis (SPOS) was pursued to synthesize short modified oligonucleotides. The standard monomer units needed for SPOS are 4,4'-dimethoxytrityl protected phosphoramidites. Herein is described the synthetic route to achieve protected DAN and NDI phosphoramidites that can be used in SPOS.

2.1.2. Scientific Question

The question this chapter attempts to answer is: *Can we design and synthesize DAN and NDI phosphoramidite derivatives that can be chemically incorporated into an oligonucleotide strand capable of adopting a favorable face-centered stacking geometry?* The successful design and synthesis of these monomers would enable facile automated solid phase oligonucleotide synthesis of oligomeric DAN and NDI or modified hybrid oligonucleotide strands with natural DNA bases.

2.1.3. Approach

In order for DAN and NDI monomers to be successfully incorporated into an oligonucleotide strand in productive fashion, certain criteria were set forth. First, the monomer had to provide the needed flexibility for DAN and NDI to adopt their preferred face-centered stacking orientation within a duplex structure. Second, in order to be chemically compatible with automated SPOS, they need to have both a phosphoramidite functional group as well as a 4,4'-dimethoxytrityl protecting group. Thirdly, the synthetic route must be reasonably accessible (reasonable yields). Fourth, the synthesis must produce enantiomeric purity suitable for biological applications.

2.1.4. Results

Monomers **2.1** (DAN) and **2.2** (NDI) were successfully synthesized with high enantiomeric excess. While the previously proposed synthetic scheme (starting with Hoffer's sugar) did not yield the desired products in reasonable yields nor stereochemical purity, redesigned monomers based upon the simplified backbone structure of glycol nucleic acids (GNA) proved to be a more synthetically achievable route. The final designs and synthetic routes were successful enough for incorporation into automated SPOS as described in Chapter 3 and Chapter 4.

2.2. BACKGROUND

The Iverson lab aromatic synthetic foldamers that are based on complementary electrostatic interactions of the aromatic units utilized flexible peptide-linked units of DAN and NDI. For the original foldamers, DAN and NDI monomers were required to have protected amino acid functional groups that were compatible with conditions used in standard solid phase peptide synthesis (SPPS). Application to oligonucleotides calls for

an entirely new monomer design of NDI and DAN phosphoramidite derivatives suitable for SPOS.

The process of SPOS involves successive iterations of removing the 4,4'-dimethoxytrityl protecting group from the 5' hydroxyl (detritylation), coupling with the next phosphoramidite monomer, capping of any unreacted hydroxyl groups with acetic anhydride and finally oxidation of the phosphotriester (P(III)) until the desired oligonucleotide length is obtained. Following assembly, the completed oligonucleotide is simultaneously deprotected and cleaved off of the solid support in base (Figure 2.1). It is critical that DAN and NDI monomer units are designed to be stable to the conditions involved with each of these steps, incorporate a 4,4'-dimethoxytrityl protected alcohol from which new monomers can be appended, and also contain the phosphoramidite functional group.

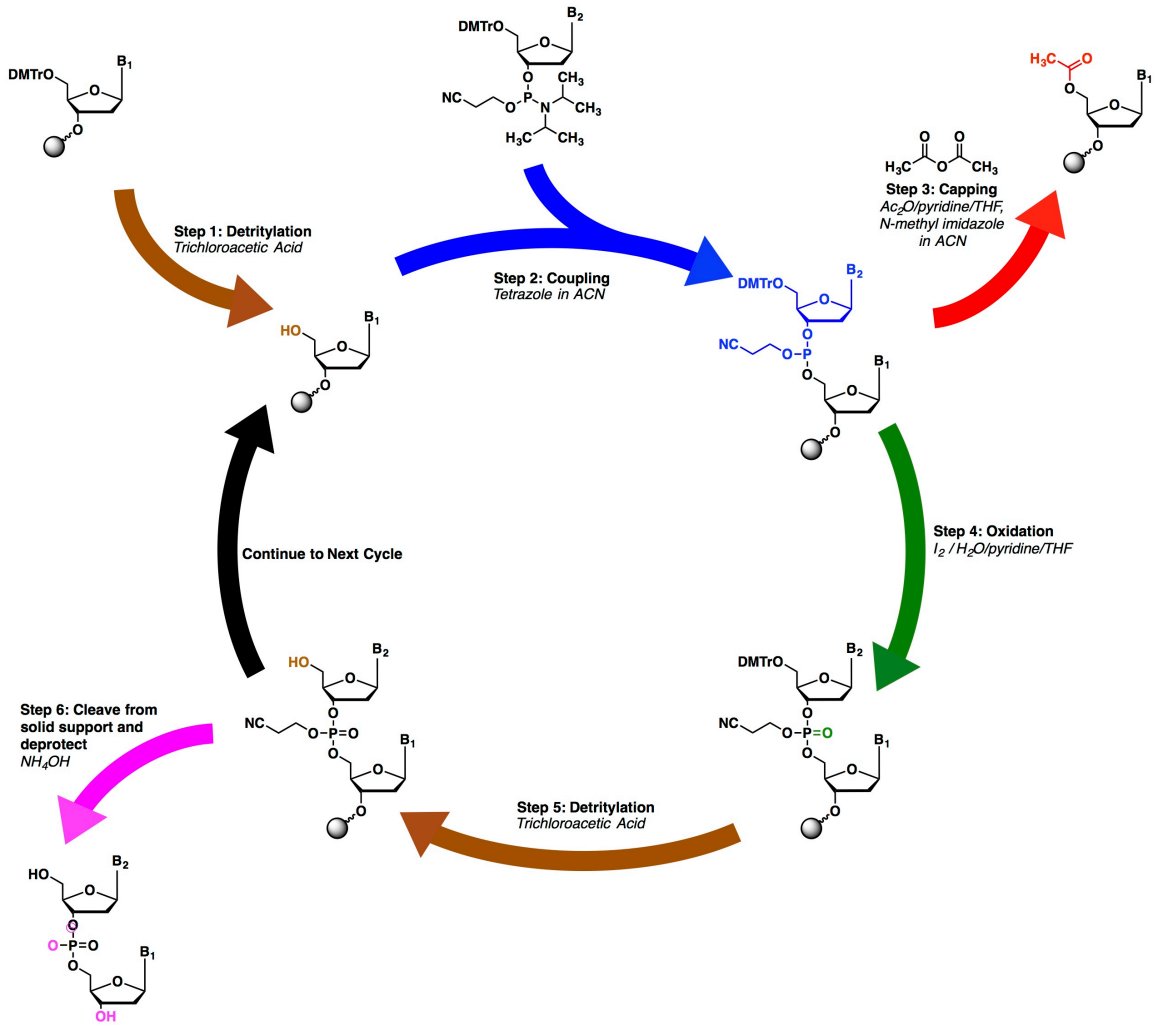


Figure 2.1 Solid phase oligonucleotide synthesis cycle. The colored arrows correspond to the resulting colored functional group transformation.

Original designs of the NDI and DAN phosphoramidite monomers consisted of the synthesis of NDI and DAN alcohols that could be used to attack Hoffer's sugar (a toluoyl protected α-chlorosugar) (Figure 2.2). However, this route proved to give unacceptably low yields that could not be overcome. Synthetic challenges with the solubility of the NDI intermediates as well as the decomposition of NDI in strongly basic

conditions proved to be insurmountable despite numerous attempts. Additionally, more complications arose due to the diasteromeric products at the anomeric position for both DAN and NDI monomers, possibly due to the facile epimerization of the acetal functional group in even slightly acidic/basic aqueous conditions. An alternative synthetic route was needed.

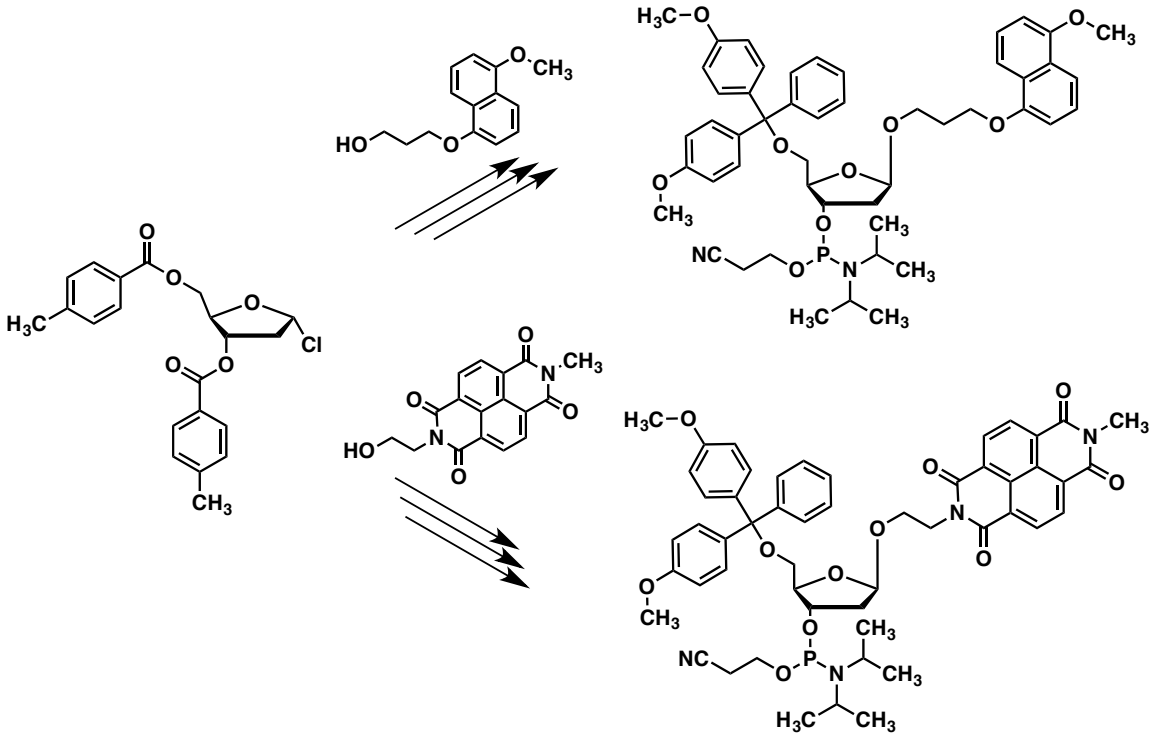


Figure 2.2 General scheme of first proposed synthesis of DAN and NDI phosphoramidite monomers starting with Hoffer's sugar, the standard starting material for modified nucleosides.

A synthetically more accessible phosphate backbone deemed glycol nucleic acids (GNA) made from a simple glycol unit has attracted increased interest over the past few decades. This simplified glycol phosphate backbone is significantly more floppy, lacking

the cyclic ribose units while also maintaining a stereochemical designation with one chiral center (Figure 2.3).

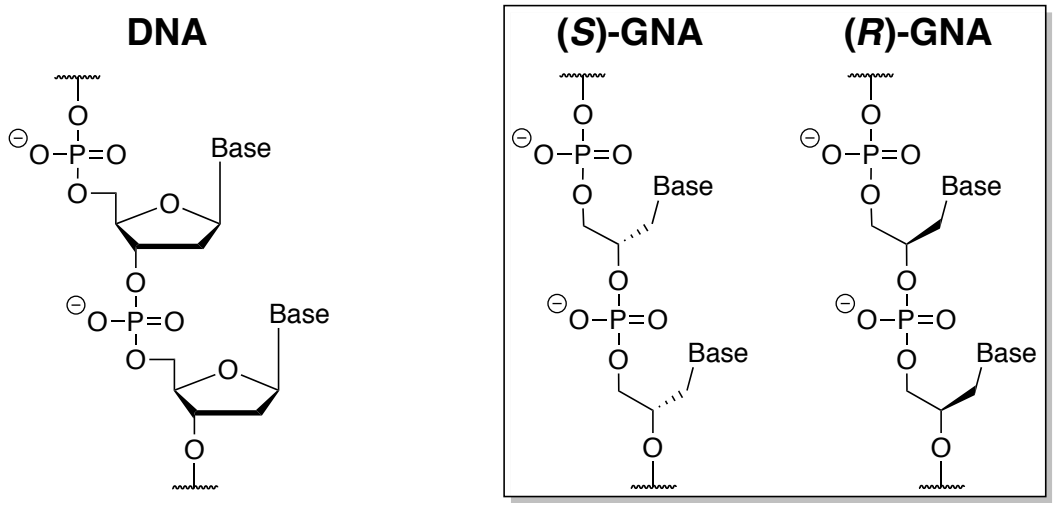
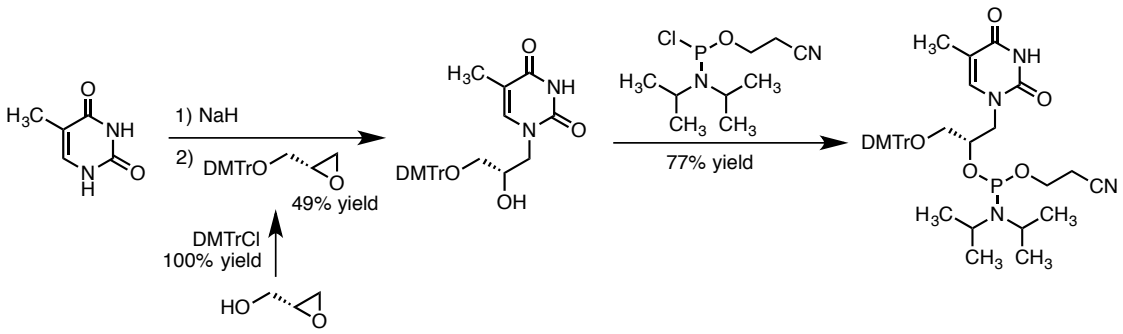


Figure 2.3 Structural comparison of natural DNA backbone with *S* and *R* GNA backbones.

Meggers and coworkers have been vigorously studying the structural, thermodynamic and biochemical properties of this modified backbone structure.[78-82] It has been demonstrated that GNA can form stable duplexes with itself. Additionally, the *S*-enantiomer has been shown to crosspair with RNA strands that do not contain G-C base pairs. More interestingly, for our present study, was that appropriately protected phosphoramidite derivatives compatible with SPOS could be synthesized in as few as 3 steps with relatively high yields (Scheme 2.1).



Scheme 2.1 Representative monomer synthesis of thymine GNA analogue phosphoramidite.[83]

Although this backbone is considerably more floppy, this flexibility could be considered an advantage, allowing the DAN and NDI modified nucleobases to adopt their preferred face-centered stacking geometry in the context of a DNA duplex.

2.3. RESULTS

2.3.1. Molecular Modeling and Design of Phosphoramidite Monomers

Preliminary molecular modeling studies were performed to give insight into the success of the initial design (utilization of the natural ribose phosphate backbone) as well as potential next-generation designs (including the utilization of the GNA backbone). Computer modeling was conducted with HyperChem 8 (HyperCube, Inc., 2007) software using the molecular mechanics AMBER force field. A model was constructed that replaced the central three base pairs of a heptamer DNA duplex with an NDI-DAN-NDI sequence such that the NDI-DAN-NDI sequence was flanked by two A-T base pairs (Figure 2.4). The conformation of the naphthyl units were taken from the structural coordinates of the previously determined crystal structure of alternating stacked DAN and NDI units.[84] The DAN and NDI units were then appended to the ribose phosphate

backbone by methylene units. The DAN unit had one more methylene unit than the NDI unit to conserve the number of atoms separating the backbone to the aromatic naphthyl core. Two designs, a short linker (m=2, n=3) and a long linker (m=3, n=4) were modeled and compared. The naphthyl units were paired with an abasic site, to keep strands in register.

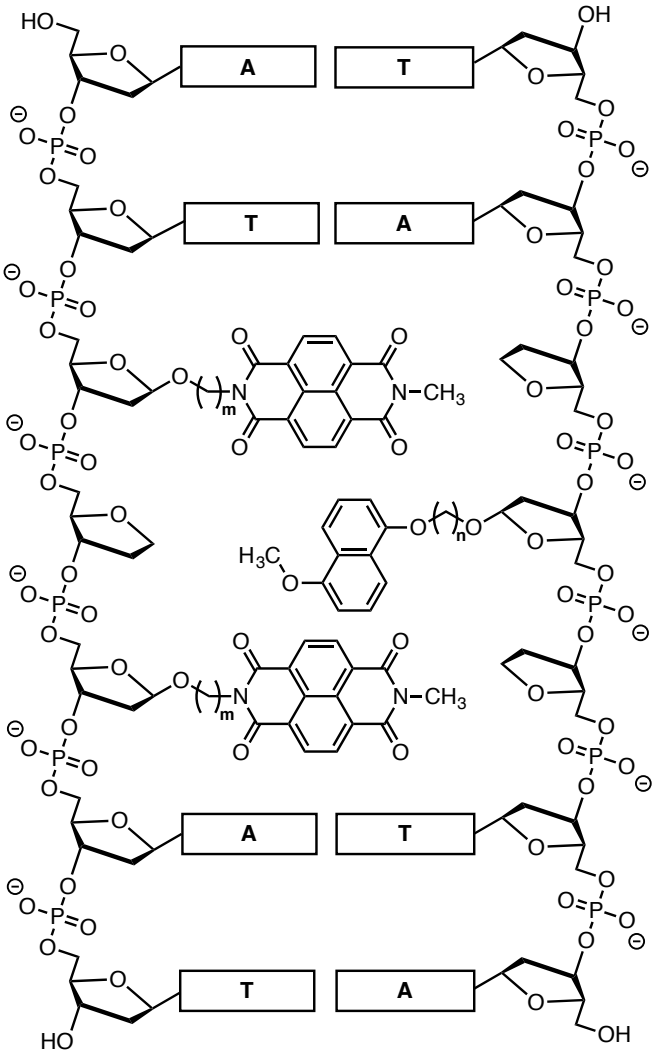
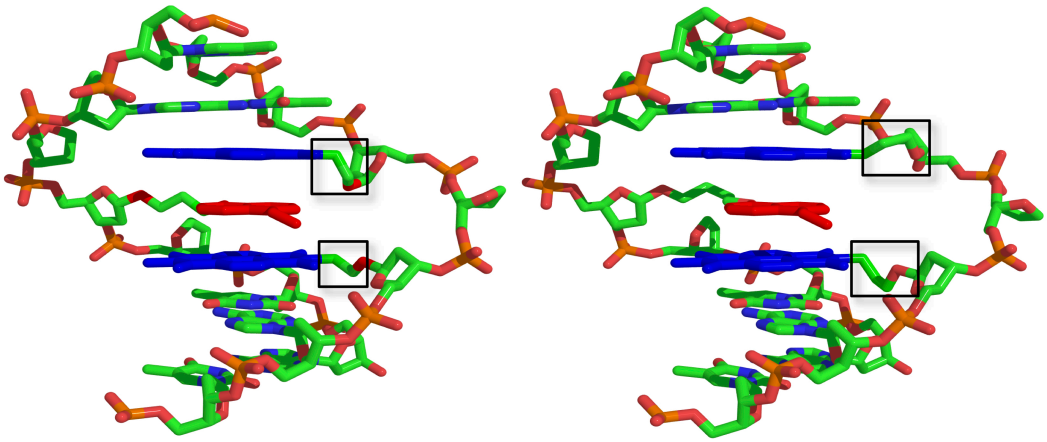


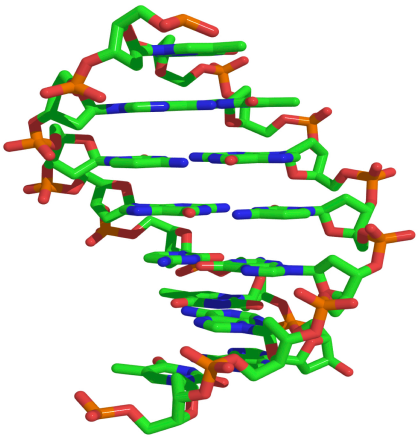
Figure 2.4 Molecular structure design of molecular modeling studies where short linker (m=2, n=3) and long linker (n=3, m=4) were studied.

The molecular dynamic calculations were carried out on the configurations of the backbone, *in vacuo*, while constraining the positions of the naphthyl units as well as the DNA bases. For both the long and short linker, 15 unique heteroduplex structures were generated. Each of these structures was subjected to geometry optimizations using a Fletcher-Reeves conjugate gradient. The resulting lowest energy structures were then visually inspected and compared to a control oligonucleotide of all natural DNA bases (Figure 2.5). By visual inspection, for both structures the backbone and linkers were able to accommodate the preferred face-centered stacking interaction of DAN and NDI. The added abasic sites imparted extra flexibility to help accommodate for the naphthyl unit stacking geometry when compared to the natural DNA duplex. However, for the long linker, there was indication of excess length in the linker as highlighted in the boxes in Figure 2.5. Therefore, it was determined that a linker length of 2 methylene units for NDI and 3 methylene units for DAN would be optimal linker lengths for the designed duplex.



Short Linker
m=2, n=3

Long Linker
m=3, n=4



5’-TAGCGTA-3’
3’-ATCGCAT-5’

Figure 2.5 Lowest energy molecular modeling results of short linker and long linker models compared to a natural DNA sequence. Boxes indicate notable differences between the two structures.

Due to the synthetic challenges of constructing a DAN and NDI ribose phosphoramidite, a new model utilizing the GNA backbone was examined utilizing the optimal linker length from the previous modeling studies (Figure 2.6). Instead of the abasic ribose sugar as a "spacer" unit, a methyl GNA unit was utilized to pair with the DAN and NDI units. After geometry optimization of the constructed model using a Fletcher-Reeves conjugate gradient, the resulting lowest energy conformation was visually inspected. This model also demonstrated that the more flexible GNA backbone was similarly able to accommodate for the preferred stacked geometry of DAN and NDI. Therefore, the phosphoramidite monomers **2.1** and **2.2** were designed (Figure 2.7) using a flexible GNA backbone linker to maintain a favorable DAN-NDI stacking geometry.

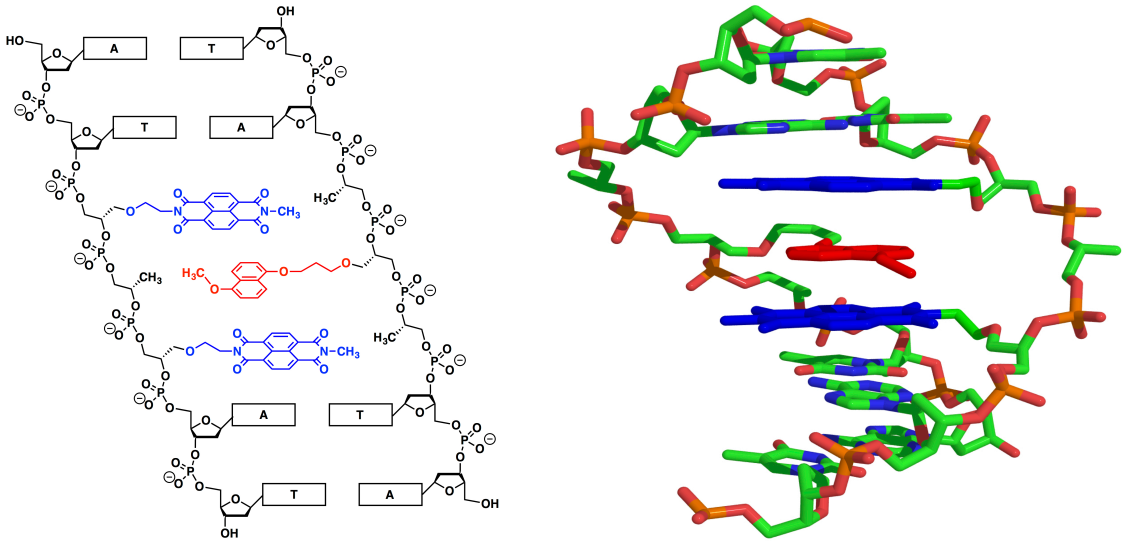
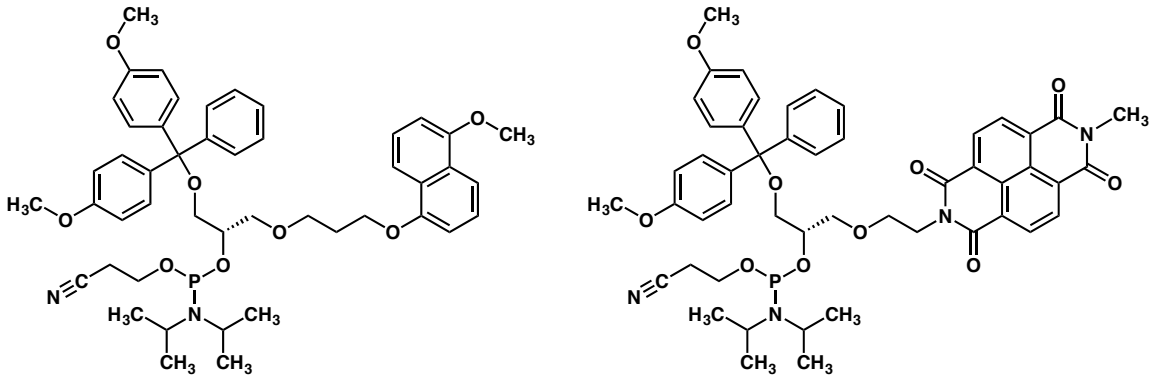


Figure 2.6 (*left*) Chemical structure and design of NDI-DAN-NDI stacking arrangement within an oligonucleotide. (*right*) Model of NDI-DAN-NDI within oligonucleotide utilizing the GNA backbone.

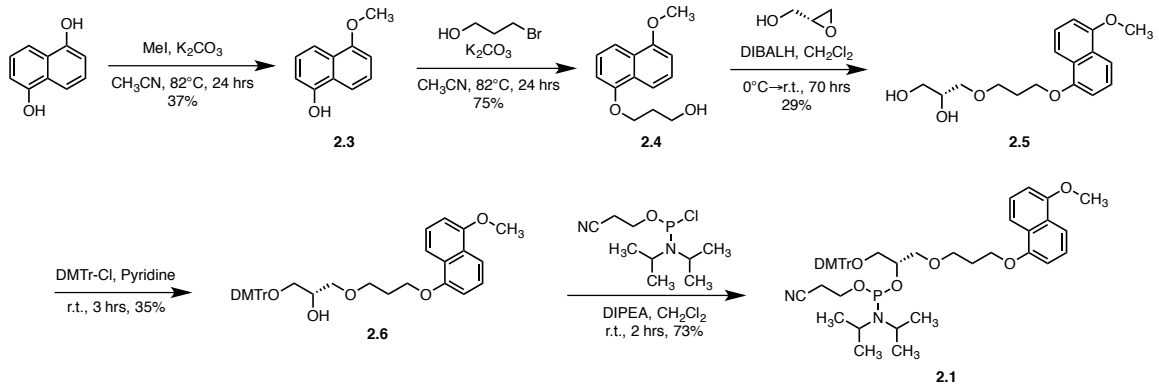


2.1 **2.2**

Figure 2.7 Phosphoramidite monomers (*left*) DAN and (*right*) NDI.

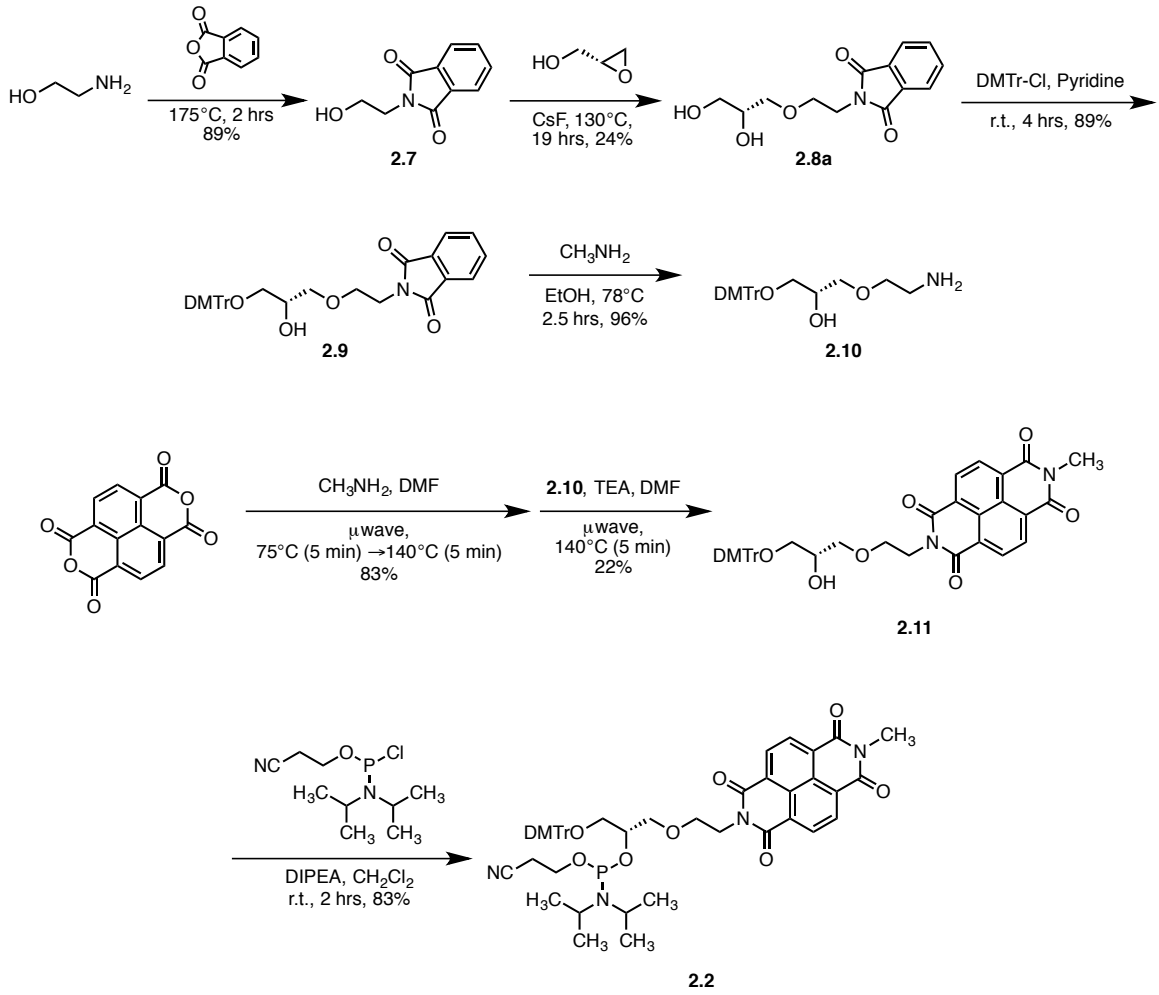
2.3.2. Synthesis of Phosphoramidites

The DAN phosphoramidite **2.1** was synthesized starting with the monomethylation of 1,5-dihydroxynaphthalene utilizing a stoichiometric amount of methyl iodide (Scheme 2.2). The product was further alkylated with 3-bromopropan-1-ol to yield alcohol **2.4**. Intermediate **2.4** was then reacted with *R*-(+)-glycidol through a regio- and enantio- specific epoxide opening mediated by DIBALH in 29% yield with an enantiomeric excess of 99% (chiral HPLC).[85] The resulting primary alcohol **2.5** was protected using 4,4'-dimethoxytrityl chloride, and subsequently transformed into the phosphoramidite **2.1** using standard conditions.[83]



Scheme 2.2 Reaction scheme to DAN monomer.

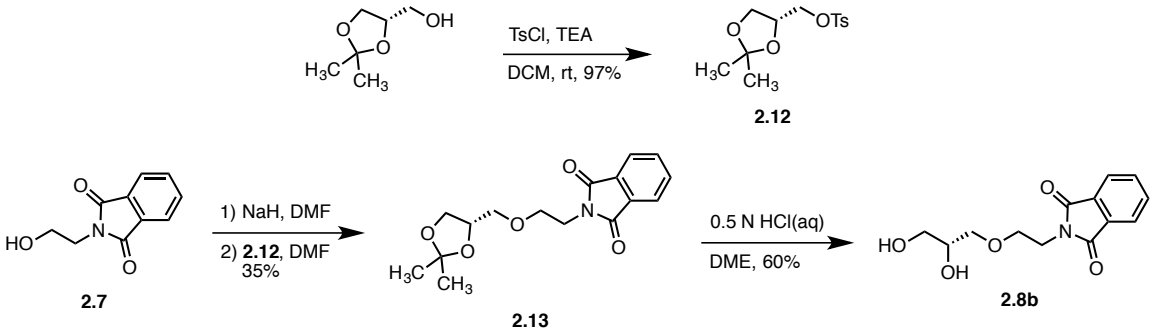
The synthesis of the NDI phosphoramidite **2.2** could not undergo a similar epoxide-opening step because the imide carbonyls of NDI are sensitive to DIBALH, so an alternative route was necessary (Scheme 2.3). This route involved a phthalimide protected ethanolamine **2.7**, which was used to open the *R*-(+)-glycidol mediated by CsF in 24% yield with an enantiomeric excess of 89% (chiral HPLC).[86] Because this is an initial study of the use of our new surrogate bases within this oligonucleotide system, the synthesis was carried on without the separation of the enantiomers. The resulting diol **2.8a** was selectively protected on the primary hydroxyl group using 4,4'-dimethoxytrityl chloride under standard conditions. The phthalimide protecting group was then removed using methylamine to yield the free primary amine **2.10**. A microwave procedure was used to append the two primary amines to 1,4,5,8-tetracarboxylic acid dianhydride: first using methylamine to obtain the methyl monoimide, and then using the free amine of **2.10** to achieve the asymmetric NDI intermediate **2.11**.[87] Intermediate **2.11** was converted to the protected phosphoramidite **2.2** using standard conditions.



Scheme 2.3 Reaction scheme to NDI monomer.

While this first synthesis proved to be successful at reaching the desired protected phosphoramidite, for biological studies, a higher enantiomeric excess would be necessary for the NDI phosphoramidite. However, for the initial studies described in Chapter 3, phosphoramidites derived from the above synthetic route were used. In order to achieve higher enantiomeric excess for the NDI monomer, an alternative synthetic route that avoids the epoxide opening step was then proposed (Scheme 2.4). This alternative route

utilizes a commercially available enantiomerically pure starting material, (*S*)-(2,2-dimethyl-1,3-dioxolan-4-yl)methanol.



Scheme 2.4 Alternate scheme to synthesize **2.8b** in higher enantiomeric excess.

Several previous syntheses utilize this enantiomerically pure starting material with strongly basic conditions without reports of altering enantiopurity.[88-89] Therefore, through this alternative synthetic route a higher enantomeric excess can be achieved without a significant reduction in yield. This synthesis required the deprotonation of the phthalimide protected ethanolamine with NaH followed by the displacement of the tosylate of **2.12**. The resulting acetal protected vicinal diol was then deprotected under aqueous acidic conditions to yield intermediate **2.8b** with an enantiomeric excess of 95% (chiral HPLC)**.** The subsequent steps to obtaining the NDI monomer were followed as from intermediate **2.8a** in Scheme 2.3.

2.4. DISCUSSION

The synthetic route to the protected phosphoramidites had to be modified from the traditional synthetic route presented by Meggers *et al*.[83] because attempts to open the protected epoxide proved to be very difficult yielding little to no product. Fortunately, the alternative approach utilizing a DIBALH mediated epoxide opening step proved to be

significantly higher yielding and with high enantiomeric excess. It was proposed that the regio- and stereo- selectivity is due to the complexation of glycidol in a four-coordinate aluminum intermediate, which activates the epoxide for nucleophilic attack at the least hindered position while preventing racemization at the central carbon position. This proved to be a successful route to obtaining the DAN phosphoramidite utilizing standard conditions to protect the primary alcohol and subsequently synthesize the phosphoramidite **2.2**.

The NDI derivative, being sensitive to DIBALH, was not synthesized using the same reaction conditions. Additionally, learning from the previously unsuccessful attempts of synthesizing the NDI monomers, the synthetic route to achieve the NDI monomer was altered to arrive at the diimide moiety only after reactions involving strong bases or strong nucleophiles. Therefore a phthalimide protected ethanolamine was synthesized in high yields with standard protocol and moved forward in this synthesis. The epoxide opening reaction with the phthalimide protected ethanolamine was carried out neat, mediated by CsF. The so-called "Cesium effect" proved to be the most successful epoxide opening reaction conditions, albeit in only modest yield (24% yield).[90] The lower enantiomeric excess of 89% gave reason to pursue a different route that could achieve an enantiomeric excess suitable for biological applications. The use of enantiomerically pure (*S*)-(2,2-dimethyl-1,3-dioxolan-4-yl)methanol enabled a synthetic route that avoided the epimerization at the single chiral center. This new synthetic route took an additional step of acetal deprotection to reveal the vicinal diol intermediate **2.8b**, but in high enantiomeric excess and with relatively high yields suitable for standard SPOS.

2.5. Conclusion

The synthetic routes for the modified protected phosphoramidites **2.1** and **2.2** were successfully achieved with high enantiomeric excess and in modest yields. Note that we had to deviate from the simple GNA monomer synthesis that had been previously well-developed. The new synthetic route utilizes a DIBALH mediated epoxide opening to get to the DAN monomer. The NDI monomer utilizes a completely different starting material, which is able to conserve its absolute stereochemistry throughout the synthesis of the NDI monomer. These two phosphoramidite monomers could then be used with automatic SPOS to synthesize oligomers of DAN and NDI along a GNA backbone or incorporate DAN and NDI within an oligonucleotide strand. The synthesis of these two monomers opened the door for new DAN- and NDI-based architectures based on the phosphodiester backbone of DNA rather than our previous foldamers that utilized peptide linkages.

2.6. Experimental Methods

5-methoxynaphthalen-1-ol (2.3). In a 100-mL round bottom flask with a stirbar, 1,5-dihydroxy naphthalene (1.00 g, 6.2 mmol) was dissolved in 32 mL acetonitrile. To the reaction mixture, K_2CO_3 (0.95 g, 6.8 mmol) and CH_3I (0.39 mL, 6.2 mmol) were added. A reflux condenser was fitted to the reaction flask and the reaction vessel was purged with argon. The reaction was stirred at reflux overnight. The reaction was cooled to room temperature and the acetonitrile was removed in *vacuo*. The thick black reaction mixture was then dissolved in $CHCl_3$ and filtered through celite. The filtrate was washed 3 times with saturated $NaHCO_3$ and 3 times with brine and dried over Na_2SO_4. The $CHCl_3$ was removed in *vacuo* and the crude product was purified with silica gel column chromatography (2% acetone in $CHCl_3$) to give a tan solid (0.41 g, 2.4 mmol, 37% yield).

Mp 128-132 °C. ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 7.85 (d, *J* = 8.5 Hz, 1H), 7.74 (d, *J* = 8.5 Hz, 1H), 7.40 (t, *J* = 7.7 Hz, 1H), 7.30 (t, *J* = 7.5 Hz, 1H), 6.85 (d, *J* = 7.4 Hz, 2H), 5.23 (s, 1H), 4.00 (s, 3H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 155.3, 151.1, 126.9, 125.3, 125.1, 114.7, 113.6, 109.4, 104.4, 55.5. HRMS-CI predicted for $C_{11}H_{11}O_2^+$ $[M+H]^+$: calcd 175.0754 *m/z*, found 175.0760 *m/z*.

3-((5-methoxynaphthalen-1-yl)oxy)propan-1-ol (2.4). In a 250-mL 3-neck round bottom flask with stirbar, **2.3** (1.47 g, 8.42 mmol) was dissolved in 84 mL acetonitrile. To the reaction mixture, K_2CO_3 (1.28 g , 9.27 mmol) and 3-bromopropan-1-ol (1.32 mL, 15.10 mmol) was added. The flask was fitted with a condenser and septa and purged with argon. The reaction mixture was heated at reflux for 24 hours. The reaction was cooled to room temperature and the acetonitrile was removed in vacuo. The thick black reaction mixture was dissolved in CH_2Cl_2 and filtered through celite. The CH_2Cl_2 was removed in vacuo and the reaction mixture was purified by silica gel column chromatography (5% acetone in CH_2Cl_2) to give a light tan solid (1.47 g, 6.33 mmol, 75% yield). Mp 100-102 °C. ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 7.83 (dd, *J* = 14.0, 8.9 Hz, 2H), 7.42 – 7.33 (m, 2H), 6.86 (t, *J* = 7.7 Hz, 2H), 4.27 (t, *J* = 5.8 Hz, 2H), 3.99 (s, 3H), 3.96 (t, *J* = 6.0 Hz, 2H), 2.18 (p, *J* = 6.0 Hz, 2H), 1.83 (s, 1H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 155.2, 154.3, 126.6, 126.5, 125.2, 125.1, 114.3, 114.0, 105.5, 104.5, 65.6, 60.5, 55.5, 32.1. HRMS-CI predicted for $C_{14}H_{16}O_3^+$ $[M]^+$: calcd 232.1099 *m/z*, found 232.1101 *m/z*.

(*R*)-3-(3-((5-methoxynaphthalen-1-yl)oxy)propoxy)propane-1,2-diol (2.5). To a clean dry 50 mL round bottom flask, was added **2.4** (0.5211 g, 2.245 mmol) and CH_2Cl_2 (15 mL). The reaction was cooled to 0 °C with an ice bath and DIBALH (1.25 mL of 1.5 M in toluene, 1.875 mmol). The reaction was taken off the ice and allowed to warm to

room temperature. The reaction was stirred for 30 minutes and then *R*-(+)-glycidol (0.1 mL, 1.506 mmol) was added to the reaction dropwise. The reaction was stirred for 72 hours at room temperature. Potassium sodium tartrate (0.5525 g, 1.959 mmol) dissolved in a minimal amount of water was added to the reaction and stirred for 30 min. The reaction was then extracted 3 times with ethyl acetate and the resulting organic layers were combined and washed with water and brine. The product mixture was then dried over Na_2SO_4 and concentrated in vacuo. The product was purified by silica gel column chromatography (4% methanol in CH_2Cl_2) to give a brown solid (0.133 g, 0.434 mmol, 29% yield). Mp 48-50 °C. ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 7.84 (d, *J* = 8.3 Hz, 2H), 7.36 (dd, *J* = 14.5, 7.8 Hz, 2H), 6.82 (d, *J* = 7.6 Hz, 2H), 4.16 (t, *J* = 6.1 Hz, 2H), 3.97 (s, 3H), 3.88 – 3.81 (m, 1H), 3.71 (t, *J* = 6.3 Hz, 2H), 3.65 (dd, *J* = 11.5, 3.6 Hz, 1H), 3.56 (dd, *J* = 11.5, 6.0 Hz, 1H), 3.49 – 3.46 (m, 2H), 3.25 (s, 2H), 2.15 (p, *J* = 6.2 Hz, 2H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 155.1, 154.2, 126.5, 125.1, 114.1, 114.0, 105.4, 104.4, 72.2, 70.6, 68.2, 64.7, 63.9, 55.4, 29.4. HRMS-ESI predicted for $C_{17}H_{22}NaO_5^+$ $[M+Na]^+$: calcd 329.1359 *m/z*, found 329.1355 *m/z*. $[\alpha]_D^{22}$ -7.3 (*c* 0.50, $CHCl_3$). 99% ee. Enantiomeric excess was determined by HPLC on a Chiralcel ODH column (0.46 cm I.D. x 25 cmL, eluent, hexane:*i*-PrOH, 95:5 v/v; flow rate, 1.0 mL/min; UV at 254 nm; room temperature; see Supporting Information)

(*S*)-1-(*bis*(4-methoxyphenyl)(phenyl)methoxy)-3-(3-((5-methoxynaphthalen-1-yl)oxy)propoxy)propan-2-ol (2.6). In a clean dry 50 mL round bottom flask, **2.5** (0.4012 g, 1.3096 mmol) and DMAP (0.0160 g, 0.131 mmol) was dissolved in pyridine (13.1 mL). The atmosphere was purged with argon and stirred. To the reaction mixture 4,4'-dimethoxytrityl chloride (0.4508 g, 1.330 mmol) was slowly added. The reaction mixture was stirred overnight and the pyridine was then removed in vacuo. The product was

purified by silica gel column chromatography (1:1 hexanes/ethyl acetate with 0.1% triethylamine) to yield a light tan oil (0.2793 g, 0.459 mmol, 35% yield). ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 7.83 (d, *J* = 8.0 Hz, 2H), 7.42 (d, *J* = 7.1 Hz, 2H), 7.38 – 7.33 (m, 2H), 7.33 – 7.27 (m, 6H), 7.23 – 7.17 (m, 1H), 6.87 – 6.77 (m, 6H), 4.18 (t, *J* = 6.1 Hz, 2H), 3.99 (s, 3H), 3.98 – 3.93 (m, 1H), 3.76 (s, 6H), 3.75 – 3.71 (m, 2H), 3.56 (ddd, *J* = 16.0, 9.7, 5.3 Hz, 2H), 3.18 (dd, *J* = 5.5, 2.9 Hz, 2H), 2.42 (s, 1H), 2.17 (p, *J* = 6.2 Hz, 2H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 155.2, 154.3, 126.6, 125.1, 114.2, 114.0, 105.4, 104.4, 72.4, 70.5, 68.3, 64.8, 64.0, 55.5, 53.4, 29.6. HRMS-ESI predicted for $C_{38}H_{40}NaO_7^+$ $[M+Na]^+$: calcd 631.2666 *m/z*, found 631.2670 *m/z*.

(*S*)-1-(*bis*(4-methoxyphenyl)(phenyl)methoxy)-3-(3-((5-methoxynaphthalen-1-yl)oxy)propoxy)propan-2-yl (2-cyanoethyl) diisopropylphosphoramidite (2.1). In a dry 15 mL round bottom flask, **2.6** (0.2793 g, 0.459 mmol) was dissolved in CH_2Cl_2 (8.2 mL) and *N*,*N*-diisopropylethylamine (0.51 mL, 2.963 mmol) was added. The reaction was stirred and purged with argon and 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.2 mL, 0.918 mmol) was added dropwise. The reaction was stirred for 3 hours at room temperature. The reaction mixture was poured into saturated aq. $NaHCO_3$, washed 3 times with CH_2Cl_2 and dried over Na_2SO_4. The $CH_2Cl_2$was removed in vacuo and the product was purified by silica gel column chromatography (4:1 hexanes/ethyl acetate with 0.1% triethylamine) to yield a mixture of diastereromers as a faint yellow oil (0.270 g, 0.334 mmol, 73% yield). ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 7.86 (t, *J* = 7.5 Hz, 2H), 7.52 – 7.44 (m, 2H), 7.41 – 7.17 (m, 11H), 6.88 – 6.75 (m, 6H), 4.18 (dt, *J* = 17.1, 6.0 Hz, 3H), 4.00 (s, 3H), 3.80 (s, 2H), 3.76 (s, 6H), 3.74 – 3.69 (m, 2H), 3.69 – 3.60 (m, 3H), 3.60 – 3.48 (m, 1H), 3.30 (dd, *J* = 9.4, 4.4 Hz, 1H), 3.26 – 3.19 (m, 1H), 3.19 – 3.12 (m, 1H), 2.57 – 2.35 (m, 2H), 2.16 (tt, *J* = 12.4, 6.1 Hz, 2H), 1.20 (d, *J* = 6.8 Hz, 6H),

1.16 (d, *J* = 6.7 Hz, 3H), 1.05 (d, *J* = 6.7 Hz, 3H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 158.5, 158.3, 155.1, 154.3, 144.9, 139.4, 136.1, 136.0, 130.0, 129.0, 128.2, 128.1, 127.8, 127.7, 127.6, 127.0, 126.6, 126.5, 125.2, 125.0, 117.7, 117.6, 114.1, 114.0, 113.9, 113.1, 112.9, 105.3, 104.4, 104.3, 85.9, 85.8, 72.7, 72.5, 72.3, 72.1, 71.8, 67.9, 67.8, 64.8, 64.1, 63.9, 58.4, 58.2, 55.4, 55.1, 43.1, 43.0, 42.9, 29.7, 29.6, 24.6, 24.5, 24.4, 20.1, 20.0. ^{31}P NMR (400 MHz, $CDCl_3$, ppm) δ 149.4, 149.3. HRMS-ESI predicted for $C_{47}H_{58}N_2O_8P^+$ $[M+H]^+$: calcd 809.3925 *m/z*, found 809.3930 *m/z*.

2-(2-hydroxyethyl)isoindoline-1,3-dione (2.7). Compound **2.7** was synthesized according to a protocol previously reported in the literature.[91] Mp 118-120 °C. ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 7.85 – 7.77 (m, 2H), 7.73 – 7.66 (m, 2H), 3.91 – 3.79 (m, 4H), 2.67 (s, 1H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 168.8, 134.0, 131.9, 123.3, 60.8, 40.7. HRMS-CI predicted for $C_{10}H_{10}NO_3^+$ $[M+H]^+$: calcd 192.0655 *m/z*, found 192.0659 *m/z*.

(*R*)-2-(2-(2,3-dihydroxypropoxy)ethyl)isoindoline-1,3-dione (2.8a). To a 50 mL round bottom flask, was added CsF (0.081 g, 0.053 mmol) and **2.7** (4.7 g, 24.6 mmol). The reaction mixture was purged with argon, stirred and heated to 135°C. The reaction was stirred for 45 min and then *R*-(+)-glycidol (1.48 mL, 22.3 mmol) was added dropwise. The reaction was stirred overnight. The reaction was then cooled to room temperature and dissolved in 4% methanol in CH_2Cl_2. The product was purified by silica gel column chromatography (4% methanol in CH_2Cl_2) to yield a white solid (1.25 g, 4.70 mmol, 24% yield). Mp 59-61 °C. ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 7.84 – 7.78 (m, 2H), 7.72 – 7.66 (m, 2H), 3.87 (t, *J* = 5.5 Hz, 2H), 3.77 (dt, *J* = 9.7, 4.8 Hz, 1H), 3.73 – 3.64 (m, 2H), 3.61 (dd, *J* = 11.5, 4.1 Hz, 1H), 3.58 – 3.49 (m, 3H), 3.30 (s, 1H), 2.83 (s, 1H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 168.5, 134.0, 131.8, 123.3, 72.4, 70.4, 68.7, 63.6, 37.4. HRMS-CI predicted for $C_{13}H_{16}NO_5^+$ $[M+H]^+$: calcd 266.1023 *m/z*, found

266.1024 *m/z*. $[\alpha]_D^{24}$ -8.7 (*c* 0.73, $CHCl_3$). 89% ee. Enantiomeric excess was determined by HPLC on a Chiralcel ODH column (0.46 cm I.D. x 25 cmL, eluent, hexane:*i*-PrOH, 95:5 v/v; flow rate, 1.0 mL/min; UV at 254 nm; room temperature)

(*S*)-2-(2-(3-(*bis*(4-methoxyphenyl)(phenyl)methoxy)-2-hydroxypropoxy)ethyl)isoindoline-1,3-dione (2.9). In a dry 15 mL round bottom flask, **2.8a** (0.3526 g, 1.33 mmol) was dissolved in pyridine (6 mL). The reaction was purged with argon, stirred and 4,4'-dimethoxytrityl chloride (0.5435 g, 1.604 mmol) was slowly added. The reaction was stirred for 4 hours and the pyridine was removed in vacuo. The product was purified by silica gel column chromatography (1:1 hexanes/ethyl acetate with 0.1% triethylamine) to yield an off white oil (0.6715 g, 1.183 mmol, 89% yield). ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 7.82 (dd, *J* = 5.4, 3.1 Hz, 2H), 7.70 (dd, *J* = 5.5, 3.0 Hz, 2H), 7.41 (d, *J* = 7.2 Hz, 2H), 7.28 (dd, *J* = 14.0, 8.4 Hz, 6H), 7.19 (t, *J* = 7.2 Hz, 1H), 6.81 (d, *J* = 8.9 Hz, 4H), 3.95 – 3.90 (m, 1H), 3.88 (t, *J* = 5.7 Hz, 2H), 3.78 (s, 6H), 3.71 (td, *J* = 5.8, 2.8 Hz, 2H), 3.62 (dd, *J* = 9.6, 3.8 Hz, 1H), 3.53 (dd, *J* = 9.6, 6.7 Hz, 1H), 3.14 (qd, *J* = 9.4, 5.6 Hz, 2H), 2.65 (d, *J* = 4.6 Hz, 1H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 168.3, 158.3, 144.8, 135.9, 133.9, 131.9, 129.9, 128.0, 127.7, 126.7, 123.2, 113.0, 85.9, 72.5, 69.7, 68.3, 64.3, 55.1, 37.4. HRMS-ESI predicted for $C_{34}H_{33}NO_7Na^+$ $[M+Na]^+$: calcd 590.2149 *m/z*, found 590.2150 *m/z*.

(*S*)-1-(2-aminoethoxy)-3-(*bis*(4-methoxyphenyl)(phenyl)methoxy)propan-2-ol (2.10). In a dry 250 mL round bottom flask, **2.9** (1.0084 g, 1.7765 mmol) was dissolved in ethanol (15 mL) and CH_3NH_2 (33% in ethanol, 30 mL, 141 mmol) was added. The reaction flask was fitted with a condenser, purged with argon, and heated to reflux for 2.5 hours. The reaction was cooled to room temperature and the solvent was removed in vacuo. The product was purified by silica gel column chromatography (10% methanol in

CH_2Cl_2 with 0.1% triethylamine) to yield an off white oil (0.4859 g, 1.111 mmol, 96% yield). ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 7.43 (d, *J* = 7.3 Hz, 2H), 7.32 (d, *J* = 8.8 Hz, 4H), 7.25 (t, *J* = 7.5 Hz, 2H), 7.16 (t, *J* = 7.3 Hz, 1H), 6.80 (d, *J* = 8.9 Hz, 4H), 3.94 (td, *J* = 9.2, 5.8 Hz, 1H), 3.72 (s, 6H), 3.59 (dd, *J* = 10.0, 3.3 Hz, 1H), 3.53 – 3.41 (m, 3H), 3.30 (s, 2H), 3.15 (ddd, *J* = 21.1, 9.3, 5.7 Hz, 2H), 2.78 (t, *J* = 5.0 Hz, 2H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 158.2, 144.6, 135.8, 129.8, 127.9, 127.5, 126.5, 112.8, 85.7, 72.7, 72.0, 69.4, 64.3, 63.2, 54.9, 41.0. HRMS-ESI predicted for $C_{26}H_{31}NO_5Na^+$ $[M+Na]^+$: calcd 460.2094 *m*/*z*, found 460.2095 *m*/*z*.

(*S*)-2-(2-(3-(*bis*(4-methoxyphenyl)(phenyl)methoxy)-2-hydroxypropoxy)ethyl)-7-methylbenzo[*lmn*][3,8]phenanthroline-1,3,6,8(2*H*,7*H*)-tetraone (2.11). In a clean dry 2-necked round bottom flask 1,4,5,8-naphthalenetetracarboxylic dianhydride (2.04 g, 7.61 mmol) was suspended in DMF (40 mL) and CH_3NH_2 (2 M in THF, 3.75 mL, 7.5 mmol) was added. The reaction mixture was sonicated for 5 minutes. The reaction was stirred and heated under microwave irradiation in the open reaction vessel fitted with a reflux condenser at 75°C for 5 minutes and then 140°C for 5 minutes. The reaction temperature was monitored by an internal probe. The reaction vessel was cooled to room temperature and the solvent was removed in vacuo. The dark brown solid was suspended in acetone and added to vigorously stirring 1N HCl. The product was filtered and washed with water. The product was then dried overnight in vacuo to yield a tan solid which was not purified any further (1.74 g, 6.17 mmol, 83% crude yield). ^{1}H NMR (400 MHz, DMSO-d_6, ppm) δ 8.73 – 8.65 (m, 4H), 3.43 (s, 3H). HRMS-CI predicted for $C_{15}H_8NO_5^+$ $[M+H]^+$: calcd 282.0397 *m*/*z*, found 282.0395 *m*/*z*.

In a clean oven dried microwave reaction vessel, the resulting crude product (0.2031 g, 0.7222 mmol) and **2.10** (0.2420 g, 0.5531 mmol) were dissolved in DMF (5 mL) and triethylamine (0.08 mL) was added. The reaction vessel was sealed and sonicated for 5 minutes. The reaction mixture was stirred and heated under microwave irradiation at 140°C for 5 minutes and then allowed to cool to room temperature. The solvent was removed in vacuo and the product was purified by silica gel column chromatography (1% methanol in CH_2Cl_2 with 0.2% triethylamine) to yield a tan frothy solid (0.0834 g, 0.1191 mmol, 22% yield). ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 8.65 (d, *J* = 1.0 Hz, 4H), 7.39 (d, *J* = 7.2 Hz, 2H), 7.25 (dd, *J* = 13.4, 8.2 Hz, 6H), 7.16 (t, *J* = 7.2 Hz, 1H), 6.76 (d, *J* = 8.7 Hz, 4H), 4.42 (t, *J* = 5.7 Hz, 2H), 3.94 – 3.88 (m, 1H), 3.84 (td, *J* = 5.7, 1.8 Hz, 2H), 3.74 (s, 6H), 3.67 (dd, *J* = 9.6, 3.8 Hz, 1H), 3.60 (dd, *J* = 9.6, 6.6 Hz, 1H), 3.56 (s, 3H), 3.18 – 3.08 (m, 2H), 2.73 (d, *J* = 4.2 Hz, 1H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 162.8, 158.4, 144.8, 136.0, 131.0, 130.9, 130.0, 128.1, 127.8, 126.7, 126.5, 126.4, 126.3, 113.0, 86.0, 72.7, 69.9, 68.2, 64.4, 55.2, 39.8, 27.4. HRMS-ESI predicted for $C_{41}H_{36}N_2O_9Na^+$ $[M+Na]^+$: calcd 723.2313 *m*/*z*, found 723.2317 *m*/*z*.

(*S*)-1-(*bis*(4-methoxyphenyl)(phenyl)methoxy)-3-(2-(7-methyl-1,3,6,8-tetraoxo-7,8-dihydrobenzo[*lmn*][3,8]phenanthrolin-2(1*H*,3*H*,6*H*)-yl)ethoxy)propan-2-yl (2-cyanoethyl) diisopropylphosphoramidite (2.2). In a dry 5 mL round bottom flask, **2.11** (0.0565 g, 0.0806 mmol) was dissolved in CH_2Cl_2 (1.5 mL) and *N*,*N*-diisopropylethylamine (0.09 mL, 0.517 mmol) was added. The reaction was stirred and purged with argon and 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.05 mL, 0.2241 mmol) was added dropwise. The reaction was stirred for 2 hours at room temperature. The reaction mixture was poured into saturated aq. $NaHCO_3$, washed 3 times with CH_2Cl_2 and dried over Na_2SO_4. The CH_2Cl_2 was removed in vacuo and the

product was purified by silica gel column chromatography (1:1 hexanes/ethyl acetate with 0.1% triethylamine) to yield a mixture of diastereomers as a faint yellow oil (0.0605 g, 0.0671 mmol, 83% yield). ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 8.72 – 8.56 (m, 4H), 7.40 (t, *J* = 7.9 Hz, 2H), 7.31 – 7.19 (m, 6H), 7.15 (dd, *J* = 14.0, 7.1 Hz, 1H), 6.82 – 6.68 (m, 4H), 4.48 – 4.33 (m, 2H), 4.14 – 4.02 (m, 1H), 3.90 – 3.76 (m, 4H), 3.74 (d, *J* = 3.6 Hz, 6H), 3.71 – 3.62 (m, 3H), 3.56 (s, 3H), 3.51 – 3.42 (m, 1H), 3.23 (dd, *J* = 9.6, 4.6 Hz, 1H), 3.20 – 3.11 (m, 1H), 3.04 (dd, *J* = 9.3, 6.4 Hz, 1H), 2.64 (tq, *J* = 7.6, 3.7 Hz, 1H), 2.49 – 2.32 (m, 1H), 1.12 (dd, *J* = 6.7, 3.8 Hz, 6H), 1.08 (d, *J* = 6.8 Hz, 3H), 0.96 (d, *J* = 6.8 Hz, 3H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 162.8, 162.5, 158.2, 144.8, 136.0, 135.9, 130.8, 130.7, 129.9, 129.8, 128.1, 128.0, 127.5, 126.5, 126.4, 126.3, 126.2, 117.9, 117.6, 112.8, 85.8, 72.6, 72.5, 72.2, 72.0, 71.9, 67.8, 64.0, 63.8, 58.4, 58.3, 58.2, 58.1, 55.0, 43.0, 42.9, 42.8, 39.7, 39.5, 27.2, 24.5, 24.4, 24.3, 24.2, 20.2, 20.1, 20.0. ^{31}P NMR (400 MHz, $CDCl_3$, ppm) δ 149.5, 149.1. HRMS-ESI predicted for $C_{50}H_{53}N_4O_{10}PNa^+$ $[M+Na]^+$: calcd 923.3392 *m/z*, found 923.3375 *m/z*.

(*R*)-(2,2-dimethyl-1,3-dioxolan-4-yl)methyl 4-methylbenzenesulfonate (2.12). In a clean, dry 50 mL round bottom flask, (*S*)-(2,2-dimethyl-1,3-dioxolan-4-yl)methanol (2 mL, 16.1924 mmol), triethylamine (3.4 mL, 24.9361 mmol) and CH_2Cl_2 (20.8 mL) were added. The reaction was stirred and purged with argon. After 15 minutes, 4-methylbenzenesulfonyl chloride (3.7044 g, 19.4308 mmol) was slowly added. The reaction was stirred at room temperature for 3 hours. The reaction was diluted with ethyl acetate (60 mL) and washed 2 times with 0.1 M HCl (aq), 2 times with saturated aq. $NaHCO_3$, and 2 times with brine. The resulting organic solution was dried over Na_2SO_4 and concentrated in vacuo. The product was purified by silica gel column chromatography (7:3 hexanes/ethyl acetate) to yield the product as a clear oil (4.5187 g,

15.7809, 97% yield). ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 7.78 (d, *J* = 8.4 Hz, 2H), 7.34 (d, *J* = 8.5 Hz, 2H), 4.30 – 4.23 (m, 1H), 4.06 – 3.93 (m, 3H), 3.75 (dd, *J* = 8.8, 5.1 Hz, 1H), 2.44 (s, 3H), 1.33 (s, 3H), 1.30 (s, 3H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 144.87, 132.37, 129.72, 127.67, 109.65, 72.70, 69.51, 65.66, 26.34, 24.88, 21.30. HRMS-ESI predicted for $C_{13}H_{18}O_5SNa^+$ $[M+Na]^+$: calcd 309.0767 *m/z*, found 309.0771 *m/z*.

(*S*)-2-(2-((2,2-dimethyl-1,3-dioxolan-4-yl)methoxy)ethyl)isoindoline-1,3-dione (2.13). In a clean, dry 25 mL round bottom flask, NaH (60% dispersion in mineral oil) (0.3381 g, 8.4514 mmol) was added and reaction vessel was purged with argon. After 15 minutes of purging with argon, DMF (6 mL) was added at room temperature. After 5 minutes, alcohol **2.7** (1.4689 g, 7.6831 mmol) was added in one portion and stirred at room temperature for 30 min. The excess solids were rinsed into the reaction mixture from the sides of the flask with additional DMF (1.4 mL). To the stirring mixture, the tosylate **2.12** (2.0676 g, 7.2207 mmol) was added dropwise. The reaction was heated to 80 °C and stirred for 2 hours. The reaction was then cooled to room temperature and diluted with ethyl acetate. The organics were washed 3 times with 1:1 brine/H_2O, 2 times with brine, dried over Na_2SO_4, and concentrated in vacuo. The product was purified by silica gel column chromatography (1:1 hexanes/ethyl acetate) to yield the product as a yellow oil (0.8242 g, 2.6994 mmol, 35% yield). ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 7.88 – 7.80 (m, 2H), 7.75 – 7.68 (m, 2H), 4.23 – 4.14 (m, 1H), 3.98 (dd, *J* = 8.2, 6.5 Hz, 1H), 3.95 – 3.83 (m, 2H), 3.81 – 3.69 (m, 2H), 3.66 (dd, *J* = 8.2, 6.5 Hz, 1H), 3.58 – 3.48 (m, 2H), 1.36 (s, 3H), 1.30 (s, 3H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 168.38, 134.09, 132.23, 123.39, 109.49, 74.77, 71.63, 68.39, 66.70, 37.30, 26.78, 25.51. HRMS-ESI predicted for $C_{16}H_{19}NO_5Na^+$ $[M+Na]^+$: calcd 328.1155 *m/z*, found 328.1161 *m/z*.

(*R*)-2-(2-(2,3-dihydroxypropoxy)ethyl)isoindoline-1,3-dione (2.8b). In a clean, dry 100 mL round bottom flask, **2.13** (0.8242 g, 2.6994 mmol) was dissolved in 1,2-dimethoxyethane (6.7 mL). To the reaction mixture, 0.5 N HCl (2.6 mL) was added. The reaction was heated to reflux and stirred for 2 hours. The reaction was allowed to cool to room temperature and concentrated in vacuo. The crude product was then diluted with ethyl acetate and washed with water, saturated aq. $NaHCO_3$ and brine. Subsequently, the organic solution was dried over Na_2SO_4 and concentrated in vacuo. The product was purified by silica gel column chromatography (ethyl acetate) to yield the product as a white solid (0.4320 g, 1.6286 mmol, 60% yield). ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 7.87 – 7.79 (m, 2H), 7.74 – 7.67 (m, 2H), 3.88 (t, *J* = 5.6 Hz, 2H), 3.82 – 3.76 (m, 1H), 3.76 – 3.67 (m, 2H), 3.63 (dd, *J* = 11.5, 4.1 Hz, 1H), 3.59 – 3.51 (m, 3H), 3.07 (s, 1H), 2.72 (s, 1H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 168.64, 134.22, 132.04, 123.49, 72.64, 70.49, 68.96, 63.82, 37.56. HRMS-ESI predicted for $C_{13}H_{15}NO_5Na^+$ $[M+Na]^+$: calcd 288.0842 *m/z*, found 288.0850 *m/z*. 95% ee. Enantiomeric excess was determined by HPLC on a Chiralcel ODH column (0.46 cm I.D. x 25 cmL, eluent, hexane:*i*-PrOH, 95:5 v/v; flow rate, 1.0 mL/min; UV at 210 nm; room temperature)

A portion of this chapter adapted with permission from Ikkanda, B.A., Samuel, S.A., and Iverson, B.L. (2014). NDI and DAN DNA: Nucleic Acid-Directed Assembly of NDI and DAN. The Journal of Organic Chemistry *79*, 2029-2037.

CHAPTER 3

Nucleic Acid-Directed Assembly of NDI and DAN

3.1. CHAPTER SUMMARY

3.1.1. Introduction

With the successful assembly of two negatively-charged chains driven by desolvation and complementary aromatic units (DAN and NDI) into an intertwined heteroduplex in water, our attention naturally turned to the context of DNA. Nucleotidomimetic foldamers consist of folded architectures of two or more oligomers, often adopting a double helical conformation in solution through combinations of intermolecular interactions. Natural DNA has a well-defined double helical architecture with incredible specificity and chain-discrimination due to Watson-Crick base-pairing, enabling it to be a scaffold for arranging the interactions of our complementary aromatic units. This chapter describes the use of natural oligonucleotides to direct the assembly of DAN and NDI sequences.

3.1.2. Scientific Questions

This chapter seeks to answer the questions: *How does the insertion of DAN and NDI monomers linked to the backbone of complementary oligonucleotides affect duplex structure and stability as compared to natural DNA bases? How does the sequence of DAN and NDI within an oligonucleotide affect structure and stability of the duplex?* Understanding these effects will give greater insight into the effect of complementary aromatic interactions on DNA duplex structure and stability as well as the particular effect of complementary aromatic units, DAN and NDI, in the context of zipper-like foldamer architectures.

3.1.3. Approach

Two novel DNA base surrogate phosphoramidites **2.1** and **2.2** (Figure 3.1), based upon the relatively electron-rich 1,5-dialkoxy-naphthalene (DAN) and relatively electron-deficient 1,4,5,8-naphthalene-tetracarboxylic diimide (NDI), respectively, were incorporated into DNA oligonucleotide strands. The NDI and DAN phosphoramidites were designed to give flexibility and spacing needed to align the NDI and DAN units in their preferred face-centered stacked conformation similar to our lab's previous aromatic foldamers. A "spacer" monomer, **3.2** (Figure 3.1), was also used to pair with the NDI or DAN unit to provide length in the backbone of the complementary DNA strand to keep the stacked aromatic residues in register.

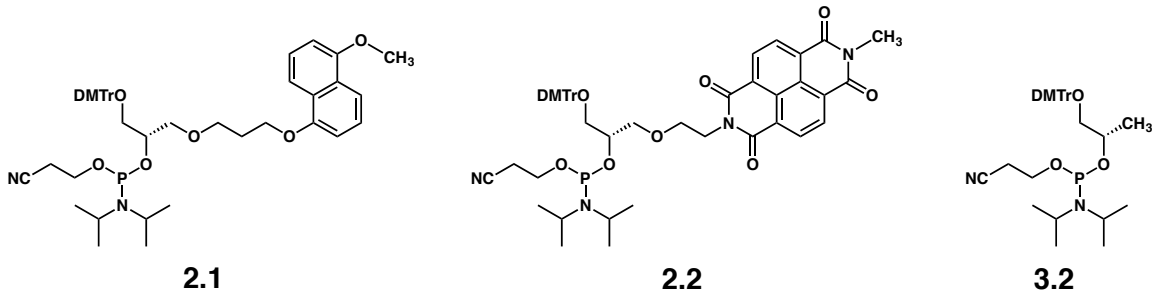


Figure 3.1 Monomer units used for oligomer synthesis.

The DAN and NDI artificial DNA bases were inserted within a 3 base pair region within the interior of a 12-mer oligonucleotide duplex in various sequential arrangements. To compare with natural DNA base pairs, a duplex with three A-T base pairs and a duplex with three G-C base pairs within the three base pair region were also synthesized. To compare with a duplex void of any aromatic bases, a duplex with three "spacer" pairs was synthesized. The structure and stability of designed and control duplexes were investigated with circular dichroism (CD) spectroscopy and UV-melting curve analysis.

3.1.4. Results

All DAN and NDI modified oligomers formed relatively stable duplexes as demonstrated by DNA thermal denaturing experiments. The CD spectra of the modified duplexes indicated double helical B-form DNA topology. Melting curve analyses revealed trends in DNA duplex stability that correlate with known association of DAN and NDI moieties in aqueous solution as well as known favorable interactions between NDI and natural DNA base-pairs. This demonstrates that DNA duplex stability and specificity can be driven by the electrostatic complementarity between DAN and NDI. In the most favorable case, an NDI-DAN-NDI arrangement in the middle of the DNA duplex was found to be approximately as stabilizing as three A-T base pairs.

3.2. BACKGROUND

3.2.1. DNA as a Scaffold to Arrange Aromatic Units

The DNA double helix is a molecular architecture endowed with numerous properties that enable it to serve as an ideal scaffold for precise arrangement of aromatic moieties. Non-covalent interactions between aromatic nucleobases, combined with desolvation effects, contribute to the DNA duplex structure, specificity, and stability (Figure 3.2). In particular, the remarkable specificity of complementary oligonucleotide strands derived from Watson–Crick hydrogen bonding can be exploited to arrange various non-natural DNA base surrogates in a highly predictable fashion. Additionally, desolvation driving forces that place the relatively hydrophobic nucleobases into the center of the double helix while placing the hydrophilic sugar phosphate backbone exposed to the aqueous environment enable the defined inter- and intra-strand stacking geometries of aromatic units.

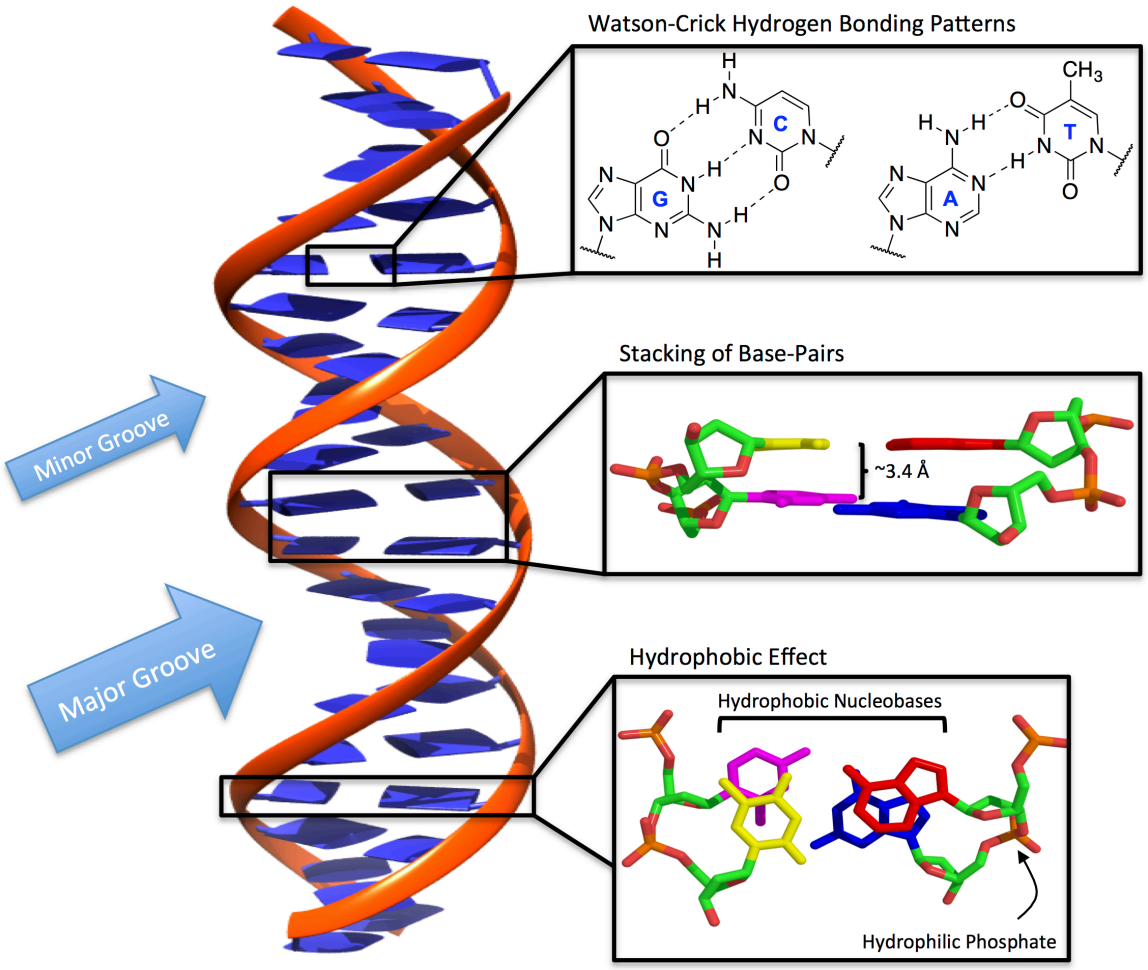


Figure 3.2 Cartoon of DNA double helix along with relevant intermolecular interactions that stabilize the double helical architecture.

Because aromatic base pairs are stacked in a predictable, ladder-like fashion, properties such as fluorescence and electron transfer as well as various complex supramolecular architectures can be investigated. Furthermore, the advanced nature of automated DNA synthesis greatly simplifies the placement of novel DNA base surrogates in a strand at any chosen location(s) within a sequence. Therefore, a number of researchers have taken on the study of modified nucleic acids that take on relevance in

the spheres of synthetic biology, nucleotidomimetic foldamers, DNA nanotechnology, fundamental biochemical understanding of nucleic acid properties, as well as even exploring the origins of life. To investigate modified nucleic acids, two general approaches have been developed: modifying the sugar phosphodiester backbone or modifying the nucleobases. The pursuit of modified backbones or modified nucleobases (or the combination thereof) has lead to tremendous success in chemical and enzymatic construction of higher order architectures.

3.2.2. Modified Backbones

A number of modified backbones have been developed over the years that have shown significant promise in intrasystem crosspairing (annealing with complementary strand of the same modified backbone) and intersystem crosspairing (annealing with complementary strands of natural DNA or RNA) (Figure 3.3). Some examples are presented here to highlight the variety of designs and biological relevance of non-natural backbones.

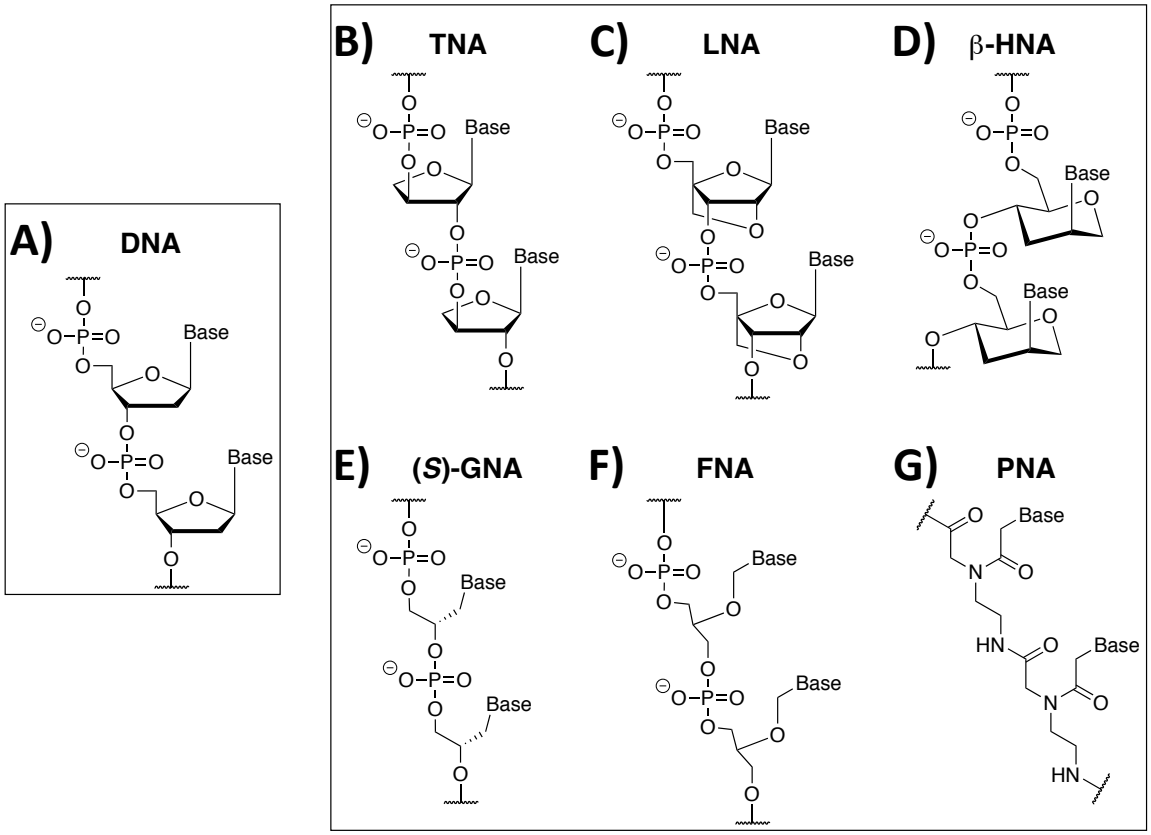


Figure 3.3 Examples of modified backbones compared to (A) natural DNA. (B) α-(L)-threofuranosyl-(3'-2') nucleic acid (C) Locked nucleic acid (D) hexitol nucleic acid (E) (*S*)-glycol nucleic acid (F) Flexible nucleic acid (G) peptide nucleic acid.

3.2.2.1. TNA: α-(L)-threofuranosyl-(3'-2') nucleic acid

Among the many nucleic acid mimics studied by Eschenmoser and coworkers, TNA has proven to be of particular interest because of its ability to form stable duplexes with complementary strands of itself as well as DNA and RNA. TNA is a simplified structure that lacks the 5' carbon of natural DNA, and alters the phosphate-furanose linkages to connect to the 2' and 3' hydroxyls of the furanose ring (Figure 3.3b).[92] Structural studies on the TNA-TNA duplex reveals that its double helical structure

resembles A-form DNA or RNA even with the slightly shorter length (by one methylene unit) between phosphate groups (Figure 3.4).[93] These not only form stable duplexes but can also be transcribed by certain polymerases with DNA templates which leads to potential biochemical applications.[94]

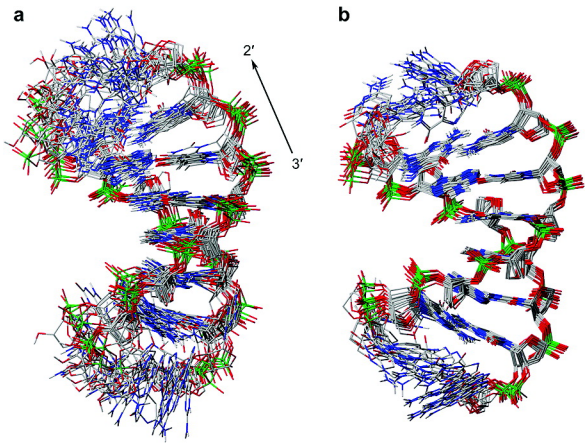


Figure 3.4 NMR structure of a TNA-TNA octamer complex in solution. Adapted with permission from Reference 93. Copyright 2008 American Chemical Society.

3.2.2.2. GNA: Glycol Nucleic Acids

Based on a simplified propylene glycol phosphate backbone, Meggers and coworkers have introduced an acyclic modified nucleic acid backbone (Figure 3.3e).[82] This greatly simplified backbone has been able to demonstrate intrasystem crosspairing as well as intersystem crosspairing with DNA. Although this backbone is significantly more floppy than the ribose phosphate backbone of natural nucleic acids, GNA has been shown to form stable duplexes based on standard Watson-Crick hydrogen bonding patterns (Figure 3.5).[79]

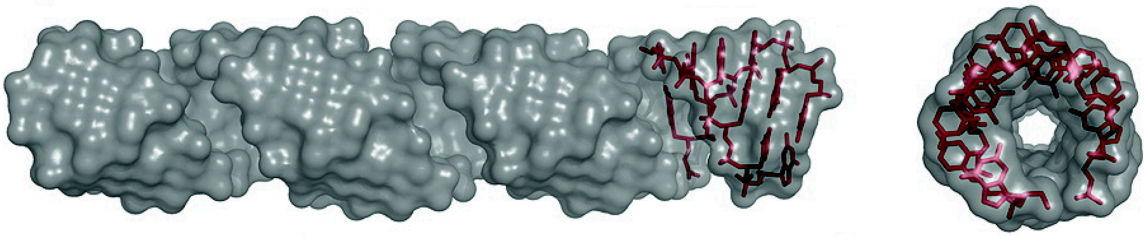


Figure 3.5 Crystal structure of GNA duplex from (*left*) side-on perspective and (*right*) top-down perspective. Adapted with permission from Reference 79. Copyright 2010 American Chemical Society.

The successful formation of higher order architecture from this very simplified backbone enables the facile introduction of novel DNA base surrogates into higher order architectures. Individual phosphoramidite monomers are readily synthesized from commercially available enantiomerically pure glycidol.

3.2.2.3. PNA: Peptide Nucleic Acids

Peptide nucleic acids are nucleotidomimetic systems that replace the sugar phosphate backbone with peptides. Several structural analogues of PNA have been developed including cyclohexyl-derived PNAs, PNA with incorporated chiral units (αPNA), or incorporation of thioester functional groups (tPNA) to name a few. However, the PNA structure that has been most widely used and studied is the 2-aminoethyl glycine backbone developed by Nielsen and Buchardt (Figure 3.3g).[95] Oligomers of PNA will crosspair in both intersystem and intrasystem fashions generating a defined double helical architecture (Figure 3.6).[96] The double helical architecture is a widened helix with a deep major groove and a shallow minor groove. This backbone has no charge and is achiral, contrasting with the sugar phosphodiester backbone in DNA, and yet adopts a well-defined helical architecture placing aromatic base pairs in the center while the peptide

backbone is on the periphery of the double helix. This particular modified backbone exhibits interesting properties such as enhanced duplex thermal stability compared to natural nucleic acid duplexes, which is thought to be derived from having a neutral backbone instead of a like negatively charged backbone that would repel its complementary strand as in natural DNA.

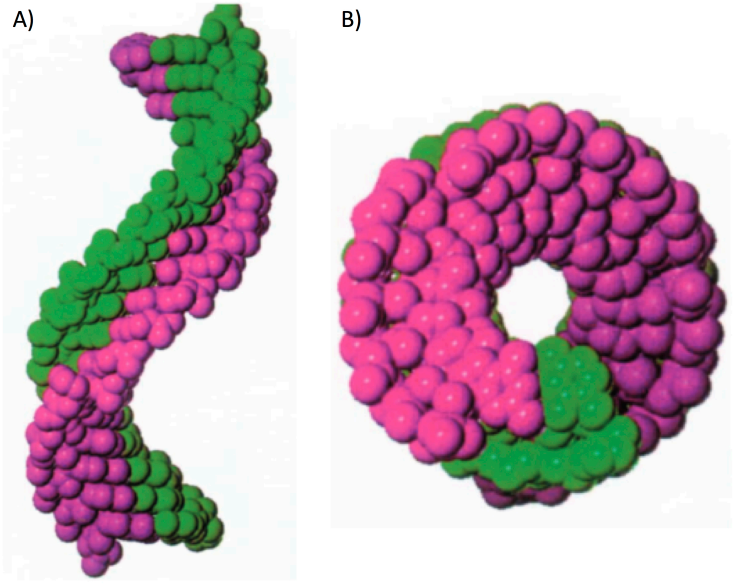


Figure 3.6 Crystal structure of PNA duplex (A) side-on view of helix. (B) top-down view of helix. Adapted by permission from Macmillan Publishers Ltd: Nature Structural and Molecular Biology[96], copyright 1997.

Oligomers of PNA have also been widely studied in the field of medicinal chemistry and biomedical research due to their ability to bind to DNA with high sequence specificity and in a variety of binding modes.[97] One of the most common binding modes of PNAs to DNA is the triplex invasion, $(PNA)_2DNA$, where a PNA-DNA duplex is formed initially through standard Watson-Crick base-pairing and the second PNA strand forms the triplex structure through Hoogsteen hydrogen bond patterns. Nielsen and

coworkers have also demonstrated other binding topologies of "tail-clamp" PNAs (Figure 3.7).[98]

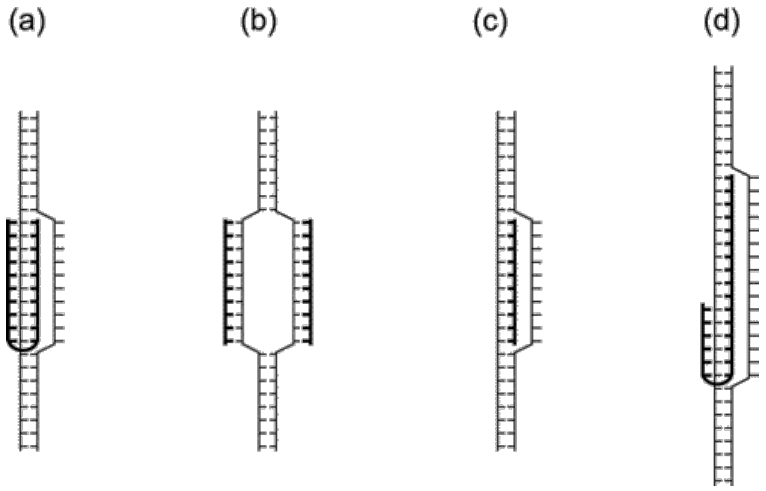


Figure 3.7 Binding modes of "Tail-clamp" PNA (a) triplex invasion complex (b) double duplex invasion complex (c) duplex invasion complex (d) principle of tail-clamp PNA invasion complex. Reprinted with permission from Reference 98. Copyright 2003 American Chemical Society.

3.2.3. Modified Bases

Numerous non-natural DNA base surrogates have been designed and synthesized that alter DNA stability and structure, including efforts to expand the genetic code with designed base pairs. Through these studies, a great deal has been learned about requirements for successful DNA duplex formation in the presence of non-natural aromatic moieties. Some examples of modified bases are discussed below.

3.2.3.1. Hydrogen Bonding Base Pairs

In efforts to expand the genetic code with designed nucleobases, a plethora of non-natural DNA bases have been designed as well as enzymatically incorporated into DNA. Through the work of Benner *et al*., designed nucleobases with alternative patterns

of hydrogen bonding have been successfully incorporated into DNA duplexes, demonstrating that the hydrogen-bonding pattern found in natural DNA bases can be altered and still produce stable duplexes (Figure 3.8).[99-100] Not only can they form stable duplexes, but it has also been demonstrated that *iso*-C and *iso*-G can be enzymatically incorporated into DNA and RNA by polymerases. Interestingly, *iso*-G has a minor tautomeric form that enables it to pair with a thymidine.

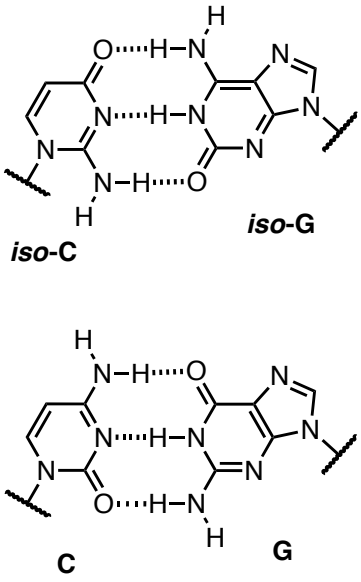


Figure 3.8 Hydrogen bonding pattern of (*top*) *iso*-cytosine (*iso*-C) and *iso*-guanosine (*iso*-G), non-natural nucleobases and (*bottom*) natural cytosine (C) and guanosine (G) nucleobases.

Extended size nucleobases (xDNA and yDNA) have been explored by Kool and coworkers (Figure 3.9). The extended size nucleobases are analogous to natural DNA bases, but have an additional benzene ring inserted extending the size of the base by about 2.4 Å.[101] Therefore, while the standard Watson-Crick hydrogen bonding patterns were designed to be conserved, the size of the nucleobase was simply extended. The resulting higher order structure is an elongated double helix with similar features to

natural B-form structure.[102-103] Additionally, all of the extended size nucleobases are fluorescent, making these useful as molecular probes. This has led to even larger expanded nucleobases and has given greater insight into the plasticity of the natural DNA backbone as well as deeper understanding of duplex structure and stability.[104]

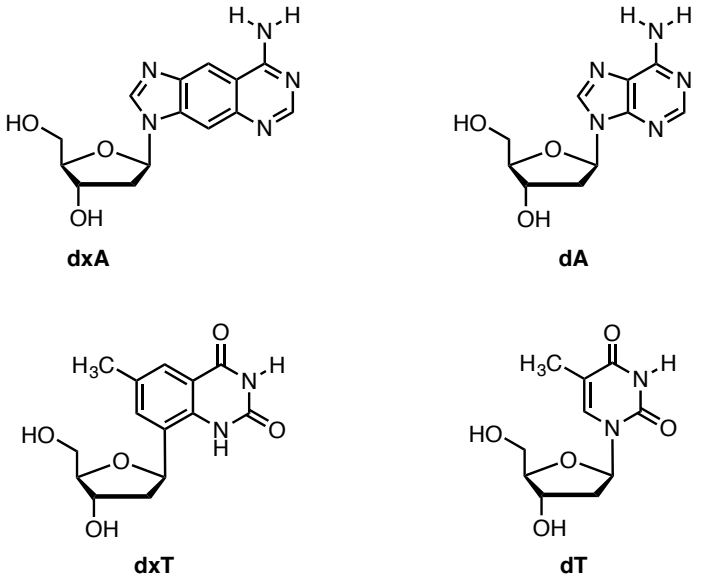


Figure 3.9 Examples of (*left*) xDNA nucleobases with inserted benzene rings compared to (*right*) their natural nucleobase analogues.

3.2.3.2. Non-Hydrogen Bonding Base Pairs

Through the synthesis and screening of an impressive number of hydrophobic, non-natural DNA base surrogates, some of which are significantly different than the structure of natural DNA bases, Romesberg and coworkers have developed several self-pair and hetero-pair nucleobases that were successfully incorporated into DNA.[105-106] In a recent study of two independent screens of 3600 candidate base pairs made from 60 non-natural nucleobases (Figure 3.10a), Romesberg examined the ability of DNA polymerase to not only recognize a non-natural base pair but extend a primer past the base pair.

Through the screening process, the pair dSICS:dMMO2 was deemed the most successful pair to be incorporated by DNA polymerases although it had a considerable limitation that the dSICS self-pair was more favorably formed than the dSICS:dMMO2 heteropair (Figure 3.10b).[107] Finally, this base pair was optimized by the observation that d5SICS had a significantly decreased rate of self-pair synthesis, encouraging the heteropairing with dMMO2 (Figure 3.10c). The wide variety of nucleobase structures that were screened provides greater insight into the requirements for successful duplex formation within the context of DNA.

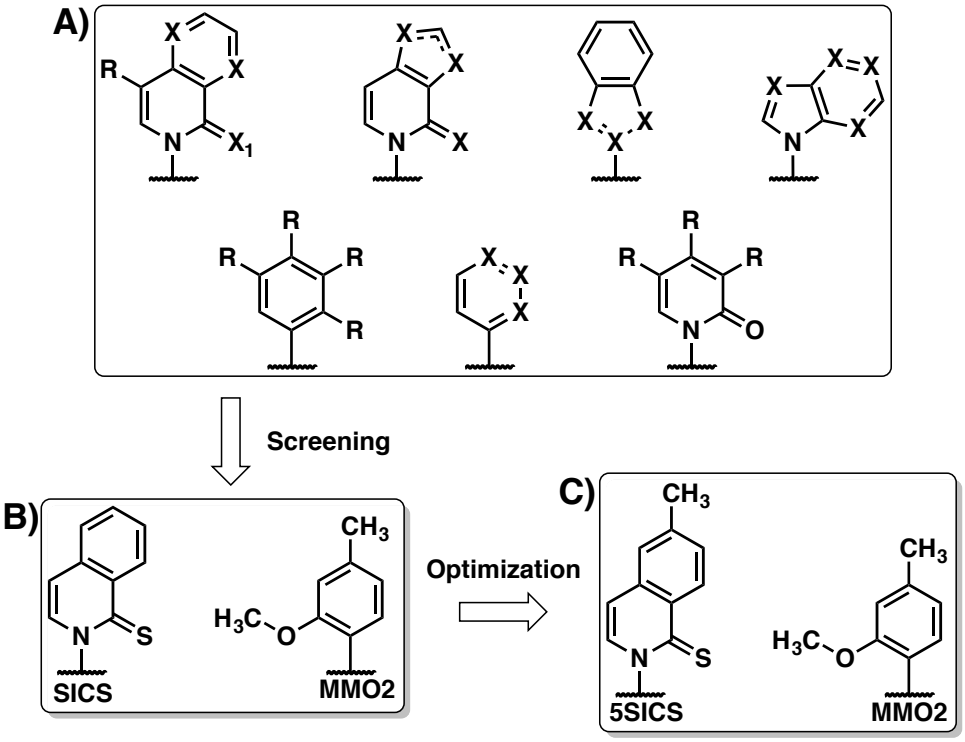


Figure 3.10 (A) Core structures of the 60 non-natural base pairs that were screened by Romesberg et al. (B) The successful base pair dSICS:dMMO2 that was chosen among numerous combinations of various base pair analogues. (C) The optimized non-natural base pair (d5SICS:dMMO2) which can be efficiently and selectively synthesized within the context of natural DNA.[107]

Kool and coworkers have also explored non-hydrogen-bonding yet isosteric analogues of DNA bases (Figure 3.11). These non-natural DNA base analogues have the closest possible structure of natural DNA bases without the hydrogen bonding capabilities as oxygen is replaced with fluorine and nitrogen with carbon while keeping the aromaticity intact.[108] It has been demonstrated that even without the hydrogen bonding ability, these isosteric nucleobases can pair with their natural counterpart as well as in their isosteric analogue pair while maintaining B-form structure.[109] These studies have provided deeper insight into the requirements of successful duplex formation as well as the role of aromatic stacking in duplex structure and stability.[110-111] Furthermore, these non-hydrogen bonding isosteres have successfully been incorporated into DNA strands by polymerases suggesting that the hydrogen-bonding patterns may not be the necessary factor for efficient and accurate enzymatic synthesis of DNA but rather, the shape of the nucleobase.[112]

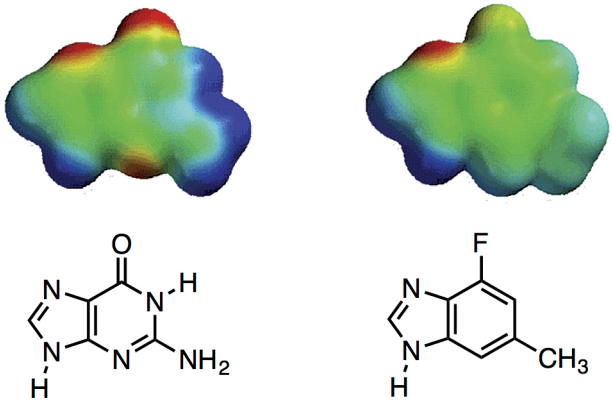


Figure 3.11 Example of (*left*) natural nucleobase guanine (G) compared with (*right*) 4-fluoro-6-methyl-1*H*-benzimidazole (H), a non-hydrogen bonding isostere. Reprinted with permission from Reference 108. Copyright 2002 American Chemical Society.

Leumann and coworkers have developed other non-natural aromatic units that promote zipper-like, stacked assembly utilizing 2'2-bipyridyl or biphenyl units.[113] These units do not utilize hydrogen-bonding patterns but rather favor interstrand stacking interactions to stabilize duplex formation. The solution structure of a DNA duplex containing a biphenyl pair reveals that B-form conformation can be maintained with the non-natural DNA bases, however with a widened helical pitch and reduced helical twist at the site of modification. The biphenyls seem to stack in a head-to-tail fashion within the duplex. Additionally, through varying the substituents of the phenyl units with electron withdrawing groups or electron donating groups, new recognition properties have been explored.[114]

3.2.3.3. Functional Organization of Chromophores

Because of its well-defined, predictable structure, DNA has also been used as a scaffold to arrange chromophores to achieve novel structures and function. For example Leumann and Grigorenko demonstrated that incorporation of a stable phenanthrenyl pair in DNA could exhibit electron transfer from an excited 5-(pyren-1-yl)uridine to 5-bromouridine as an electron acceptor (Figure 3.12).[115] The hydrophobic phenanthrenyl pair is void of hydrogen bonding capabilities and is assembled through interstrand stacking interactions. This idea was further explored as multiple phenanthrenyl base pairs were introduced and enhanced excess electron transfer was observed. These results demonstrate the potential for DNA assembled modified nucleobases as potential assemblies with tuned electronic properties.

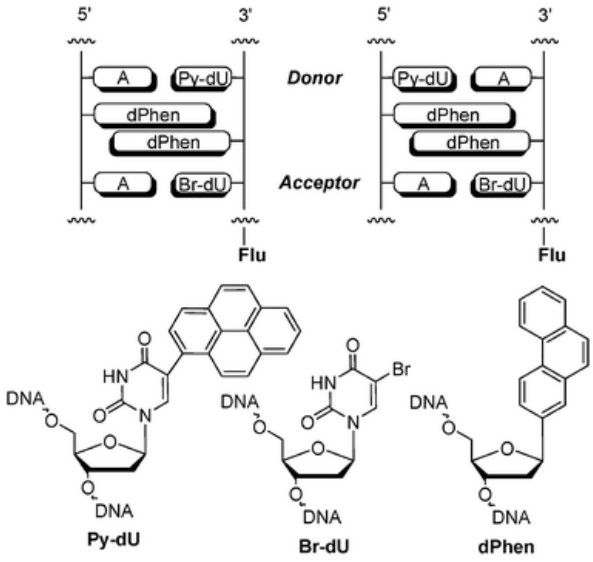


Figure 3.12 Diagram of system that displays electron transfer between a phenanthrenyl pair in DNA. Reproduced from Reference 115 with permission of The Royal Society of Chemistry.

Fluorescence properties have been introduced into DNA strands via the organization of multiple chromophores within a DNA duplex. For example, fluorescent properties were observed when multiple perylenediimide chromophores were arranged in a zipper-like conformation.[116] Through the work of Wagenknecht *et al*. as perylenediimides were introduced into a DNA strand across from an abasic site. The dimers as well as the hexamers of perylendiimide within a DNA duplex exhibit characteristic eximer-type fluorescence. These unique properties show promise as an indicator of duplex formation by their fluorescence properties upon stacking.

Additionally, Kool and co-workers have also studied a series of oligomeric polyfluorophores that can be arranged onto a DNA backbone (Figure 3.13).[117-118] Their results demonstrate that not only the character of the fluorophore dictates its properties, but also the sequence as it is arranged on the DNA backbone has an effect on their

fluorescence properties. These fluorescent nucleobase analogs along with others have had tremendous use in biotechnological applications as probes of interactions with DNA and RNA. For example, a tetramer of oligodeoxyriboside fluorophores (ODFs) has been successfully used to sense enzymatic bond-cleaving activity of esterases and lipases.[119]

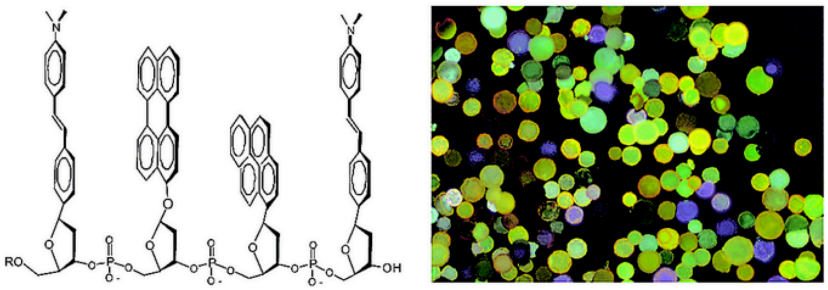


Figure 3.13 (*left*) Composite polyfluors that are synthesized from fluorescent deoxyribosides and (*right*) fluorescent microscope image of a library of polyfluors. Reprinted with permission from Reference 118. Copyright 2002 American Chemical Society.

More recently, ODFs have been used in applications of detecting toxic metal ions (Cd^{II}, Cr^{VI}, Cr^{III}, Fe^{III}, Hg^{II}, $MeHg^{II}$, Mn^{II}, and Ni^{II})[120] and anions (cyanide, selenite, thiocyanate and arsenate)[121] in water. ODFs can be easily synthesized by automated DNA synthesis techniques and could therefore be easily coupled to metal-binding deoxyribonucleotide monomers. Through various screening studies of different sequences of ODFs, Kool *et al*. has demonstrated the versatility and sensitivity of ODFs as chemosensors by a simple optical method.[117]

3.2.3.4. Higher Order DNA Assembly

Lewis et al. had developed perylenediimide modified oligonucleotides with higher order molecular structures, such as oligonucleotides that can form either duplex or hairpin dimer structures depending on the oligonucleotide sequence.[122] Additionally,

DNA duplexes appended to perylenediimides can self-assemble into supramolecular polymers in water. Perylenediimide has also been incorporated into DNA by Li and coworkers to develop thermophilic foldamers that exhibit heat promoted folding (Figure 3.14).[123] It is proposed that the higher order architecture is due to hydrophobic effects driving the perylenediimide units together.

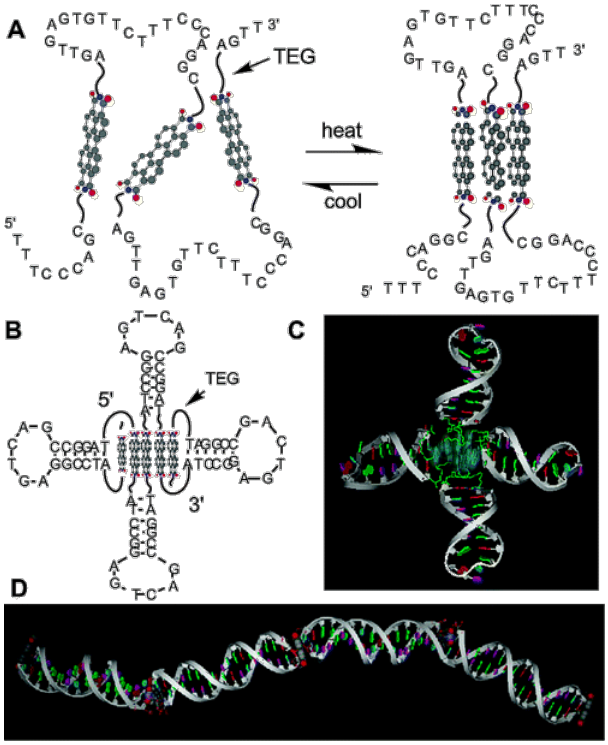


Figure 3.14 Schematic diagram of thermophilic folding of perylenediimide-linked strands of DNA to generate a higher order architecture. Reprinted with permission from Reference 123. Copyright 2003 American Chemical Society.

Majima and coworkers have successfully utilized DNA as a scaffold to direct the assembly of porphyrins (Figure 3.15). For example, a porphyrin dimer was assembled to stack on the exterior of a DNA double helix upon DNA duplex formation.[124] DNA-

porphyrin conjugates have also been assembled such that four single oligonucleotide strands are appended to a porphyrin and subsequently annealed to a second porphyrin appended to four complementary oligonucleotides. This demonstrated the use of DNA to direct even higher orders of assembly generating a scaffold for a host-guest system.[125]

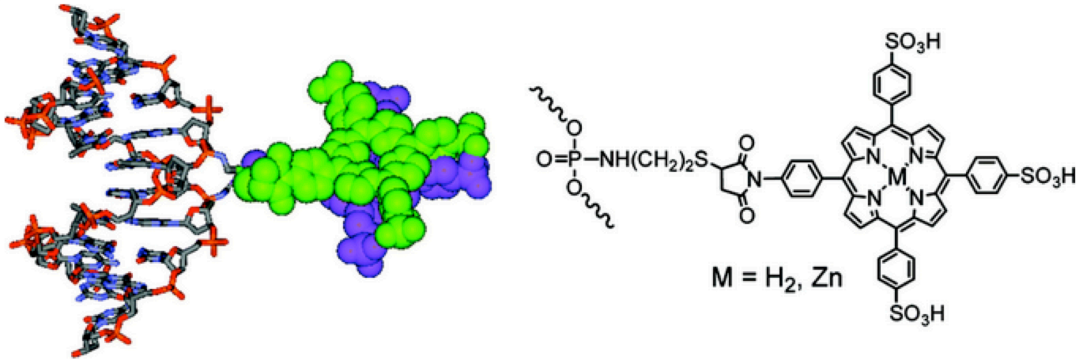


Figure 3.15 (*left*) porphyrin dimer assembled on the exterior of a DNA duplex. (*right*) molecular structure of porphyrin appended to the DNA duplex. Reprinted with permission from Reference 124. Copyright 2008 American Chemical Society.

3.2.4. NDI and DAN Interactions and Assemblies

While the previous examples utilized a variety of non-covalent interactions to drive duplex assembly, our lab has been exploring the use of relatively electron-rich 1,5-dialkoxynaphthalene (DAN) and relatively electron-deficient 1,4,5,8-naphthalenetetracarboxylic diimide (NDI) derivatives in aqueous solution to create novel folded and assembled structures based on alternating face-centered stacking of the DAN and NDI units. Importantly, in strongly interacting solvents such as water, NDI and DAN have an association constant that is 1 or 2 orders of magnitude larger than the self-association constant of either NDI or DAN, respectively. This specificity is thought to be due to complementary electrostatic interactions that can best be rationalized by focusing

on the local and direct interactions of the highly polarized substituents (i.e., the diimide carbonyl groups of NDI and the ether oxygen atoms of DAN) on the periphery of the aromatic rings. Such considerations explain why DAN and NDI adopt a fully face-centered stacking geometry in the solid state while NDI self-stacks in an offset mode and DAN does not prefer to self-stack but rather adopts a herringbone geometry in the solid state. In the context of desolvation, these preferred electrostatics-driven geometries would favor DAN–NDI association over NDI or DAN self-association in water. Interestingly, in previous work involving duplex assembly from relatively flexible amide-linked chains of DAN and NDI units, the free energy of duplex formation decreased only slightly as the temperature increased, demonstrating an apparent enthalpy–entropy compensation effect.

Herein is described the incorporation of novel DAN and NDI DNA nucleobase surrogates into DNA oligonucleotides. The stabilities and structures of various DAN- and NDI-modified oligomers were investigated. Complementary oligonucleotides that assembled to allow for alternating NDI–DAN–NDI stacking proved to be more stable than any of the other combinations investigated, demonstrating, to the best of our knowledge, for the first time that a stacking preference based on electrostatic complementarity can drive duplex stability and specificity.

3.3. RESULTS

3.3.1. Sequence Design

In comparing the interactions between base pairs in natural DNA and the face-centered stacking of DAN and NDI in the solid state, it is proposed that this observed geometry of DAN and NDI will be maintained within the context of natural DNA base pairs (Figure 3.16).

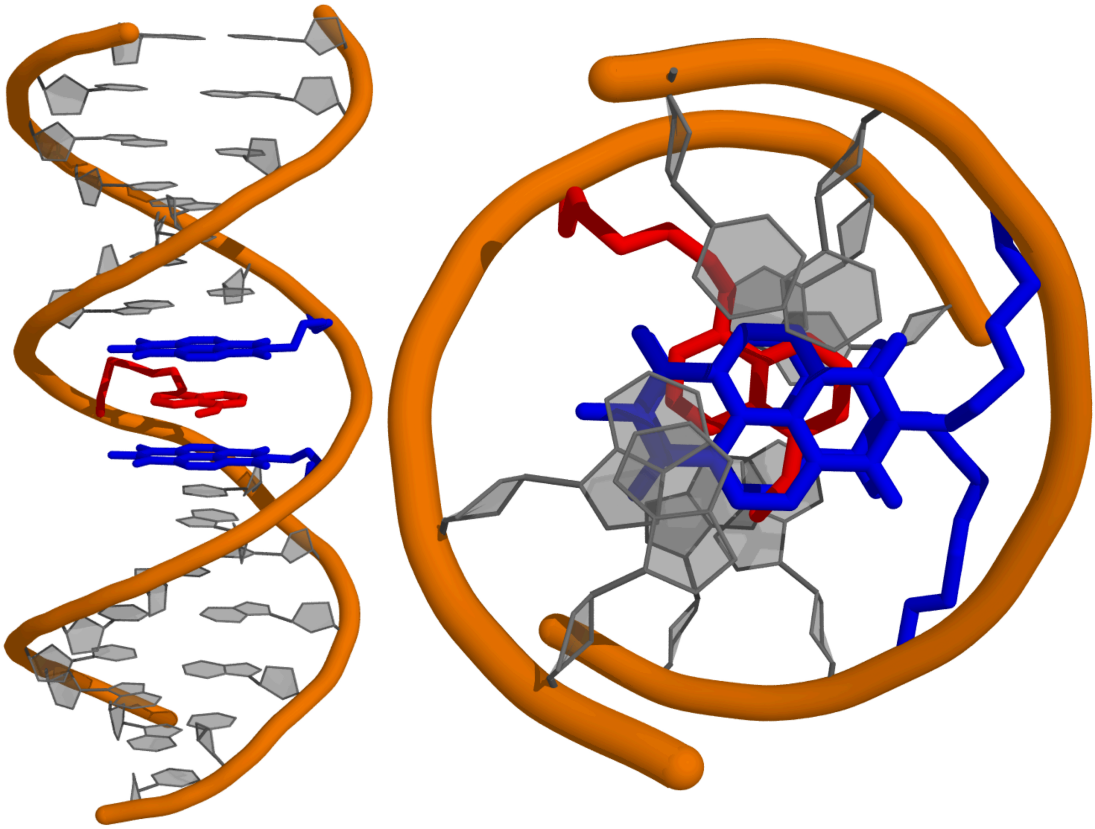


Figure 3.16 Model of the interaction of NDI and DAN appended to an oligonucleotide backbone flanked by natural DNA bases. (*left*) side-on perspective (*right*) top-down perspective.

In order to investigate the effects of incorporating DAN and NDI artificial DNA bases, a region of 3 base-pairs was inserted into the interior of **duplex 3.1**, the conserved DNA sequence of 12 base pairs (Figure 3.17). Four single-stranded 15-mer oligonucleotides with NDI and DAN modifications were designed to generate **duplexes 3.5-3.8** as well as the oligonucleotides for the three control duplexes with three A-T, G-C or spacer base pairs (**duplexes 3.2-3.4**, respectively). Control **duplex 3.2** and **duplex 3.3** both contain a deoxyribose backbone as a reference in order to compare our modified duplex to a natural DNA duplex, noting that substituting a GNA backbone within a

sequence of natural bases has been shown to destabilize the DNA duplex melting temperature.[81] **Duplex 3.5** is designed to examine the stability provided by three DAN units stacking while **duplex 3.6** is designed to examine the stability provided by alternating DAN-NDI-DAN stacking. **Duplex 3.7** is designed to examine stability provided by three NDI units stacking, while **duplex 3.8** is designed to examine stability provided by alternating NDI-DAN-NDI stacking. Since NDI is a known strong DNA intercalator[76, 126] and has been shown to have relatively high affinity for G-quadruplex DNA,[127-128] it is therefore predicted to have greater association with natural DNA bases than DAN. It follows that **duplex 3.8** (with 2 NDI-nucleobase contacts) is predicted to provide greater stability than **duplex 3.6** (with 2 DAN-nucleobase contacts). For the same reason, and additionally because NDI is known to produce relatively stable off-set self-stacks compared to DAN self-stacks, **duplex 3.7** is predicted to provide significantly greater stability than **duplex 3.5**. Taken together, analyzing the relative stabilities of **duplexes 3.1-3.8** will allow an assessment of the role that electrostatic complementarity of DAN and NDI plays in stabilizing DNA duplexes, and how such stability compares to natural A-T versus G-C sequences.

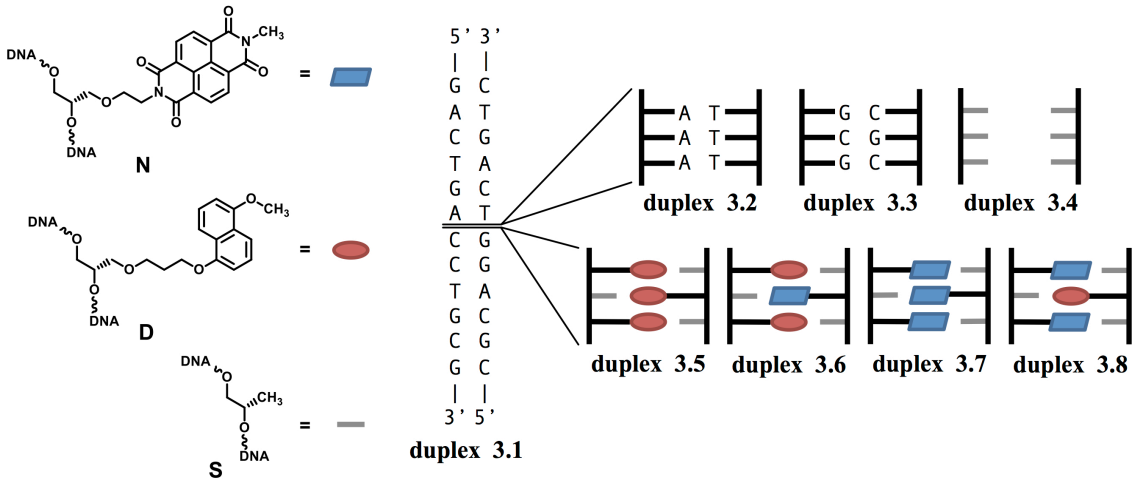
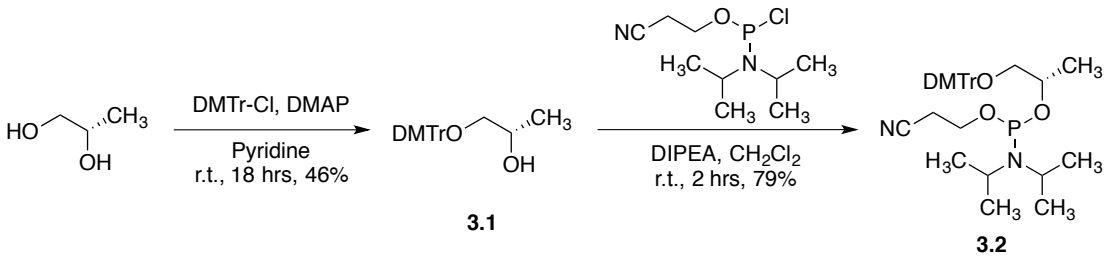


Figure 3.17 Modified DNA base surrogates and a cartoon representing the control DNA **duplex 3.1** as well as the insertions of natural DNA bases (**duplex 3.2** and **duplex 3.3**), spacer units (**duplex 3.4**), and the four NDI- and DAN-modified units (**duplexes 3.5-3.8**)

3.3.2. Synthesis of Oligonucleotides

Oligonucleotides were synthesized on an automated nucleic acid synthesizer according to standard automated oligonucleotide synthesis protocols. Monomers **2.1** and **2.2** were synthesized using the protocol developed in Chapter 2. Monomer **3.2** was synthesized using conditions similar to previous syntheses (Scheme 3.1).[129]



Scheme 3.1 Reaction scheme to obtain "spacer" phosphoramidite

The insertion of these modified building blocks required alternative conditions for the DNA synthesizer. These modified phosphoramidites were dissolved in a 3:1 CH_2Cl_2:CH_3CN solution to adequately solubilize the monomers. To avoid aminolysis of the imide functional group in NDI, oligonucleotides containing NDI were synthesized utilizing UltraMild synthesis and deprotection methods from Glen Research which avoid the concentrated aqueous ammonia cleavage step. Oligonucleotides were purified by reverse phase HPLC and characterized by HRMS-ESI (negative mode, CH_3CN/aqueous ammonium carbonate) (Table 3.1).

Sequence	Ion	Mass (calcd)	Mass (found)
5'-GACTGACCTGCG-3'	$[M-4H]^{-4}$	910.1545	910.1540
5'-CGCAGGTCAGTC-3'	$[M-4H]^{-4}$	910.1545	910.1537
5'-GACTGAAAACCTGCG-3'	$[M-5H]^{-5}$	915.7567	915.7570
5'-CGCAGGTTTTCAGTC-3'	$[M-3H]^{-3}$	1517.9211	1517.9144
5'-GACTGAGCGCCTGCG-3'	$[M-5H]^{-5}$	917.3524	917.3528
5'-CGCAGGCGCTCAGTC-3'	$[M-3H]^{-3}$	1516.2568	1516.2541
5'-GACTGASSSCCTGCG-3'	$[M-4H]^{-4}$	1013.6607	1013.6586
5'-CGCAGGSSSTCAGTC-3'	$[M-5H]^{-5}$	810.7271	810.7270
5'-GACTGADSDCCTGCG-3'	$[M-3H]^{-3}$	1505.2795	1505.2785
5'-CGCAGGSNSTCAGTC-3'	$[M-5H]^{-5}$	875.1389	875.1373
5'-GACTGANSNCCTGCG-3'	$[M-5H]^{-5}$	939.5506	939.5497
5'-CGCAGGSDSTCAGTC-3'	$[M-4H]^{-4}$	1071.1842	1071.1818

Table 3.1 HRMS-ESI data for oligonucleotides used in experiments. "S" refers to the methyl spacer unit, "N" refers to the NDI unit, and "D" refers to the DAN unit.

3.3.3. Thermal Denaturing Studies

Thermal denaturing studies were conducted to observe the effect of DAN and NDI interactions on duplex stability (Table 3.2). The highest observed melting temperature among the duplexes was observed for **duplex 3.3** containing three G-C base pairs, which was 11°C higher than that of the control **duplex 3.1**. The next most stable

duplex as **duplex 3.2** containing A-T base pairs which exhibited a 4°C increase in melting temperature compared with **duplex 3.1**. The duplex containing methyl groups instead of any natural base pairs or aromatic units was by far the least stable duplex examined (having a melting temperature 30°C lower than that of control **duplex 3.1**). Notably, replacing just three of the 6 methyl units in the region of interest with either DAN or NDI (**duplexes 3.5-3.8**) increased the melting temperature by at least 18°C.

Duplex	Sequence	T_m (°C)	ΔT_m (°C)
3.1	5'-GACTGACCTGCG-3' 3'-CTGACTGGACGC-5'	54	-
3.2	5'-GACTGA AAA CCTGCG-3' 3'-CTGACT TTT GGACGC-5'	58	4
3.3	5'-GACTGA GCG CCTGCG-3' 3'-CTGACT CGC GGACGC-5'	65	11
3.4	5'-GACTGA SSS CCTGCG-3' 3'-CTGACT SSS GGACGC-5'	24	-30
3.5	5'-GACTGA DSD CCTGCG-3' 3'-CTGACT SDS GGACGC-5'	42	-12
3.6	5'-GACTGA DSD CCTGCG-3' 3'-CTGACT SNS GGACGC-5'	52	-2
3.7	5'-GACTGA NSN CCTGCG-3' 3'-CTGACT SNS GGACGC-5'	53	-1
3.8	5'-GACTGA NSN CCTGCG-3' 3'-CTGACT SDS GGACGC-5'	57	3

Table 3.2 T_m data for the DNA duplexes. DNA melting experiments were carried out at a duplex concentration of 1.5 μM (pH 7, 100 mM NaCl, 10 mM NaH_2PO_4, 0.1 mM EDTA).

As expected, the modified duplex with the highest melting temperature was **duplex 3.8**. This duplex showed an increase in melting temperature of 3°C compared to **duplex 3.1**. Notably, this is about the same increase in melting temperature afforded by three A-T base pairs (**duplex 3.2**) but not as large as that seen with three G-C base pairs (**duplex 3.3**). Amongst the modified duplexes, there is a 4°C increase in melting temperature from **duplex 3.7** to **duplex 3.8** which is merely changing the central NDI unit to the more electrostatically complementary DAN unit resulting in an NDI-DAN-NDI arrangement. Comparing **duplex 3.5** with **duplex 3.6** demonstrates that a homo-DAN arrangement is also significantly less stabilizing than a DAN-NDI-DAN alternating arrangement, as seen in the 10°C difference. Finally, it is also observed that NDI-nucleobase interactions appeared to be preferred to DAN-nucleobase interactions as seen in the 5°C increase in melting temperature from **duplex 3.6** to **duplex 3.8**.

3.3.4. Circular Dichroism Spectroscopy

Structural evidence of duplex formation was investigated through CD spectroscopy (Figure 3.18). Each of the modified duplexes showed spectral features consistent with an overall B-form DNA topology, exhibiting a bisignate spectrum with a null occurring near λ = 260 nm. In addition, CD spectra at various temperatures ranging from 15 to 70°C demonstrated behavior similar to that of the corresponding **duplex 3.1** CD spectra. This further supports its B-form structure. Understandably, these results do not exclude the presence of any regions with substantial deviations from B-form structure in any of the duplexes as CD spectroscopy is not a high-resolution technique. Therefore, little to no insight can be gained regarding the exact stacking topologies of the modified bases in **duplexes 3.4-3.8** from the CD spectra alone. And while observation of the typical charge-transfer absorbance can provide some support of the face-centered

interaction of DAN and NDI, the concentrations of the solutions that were studied were far too small to observe any significant CD signal, let alone changes in CD spectra, resulting from the NDI absorbance in the 300-400 nm range.

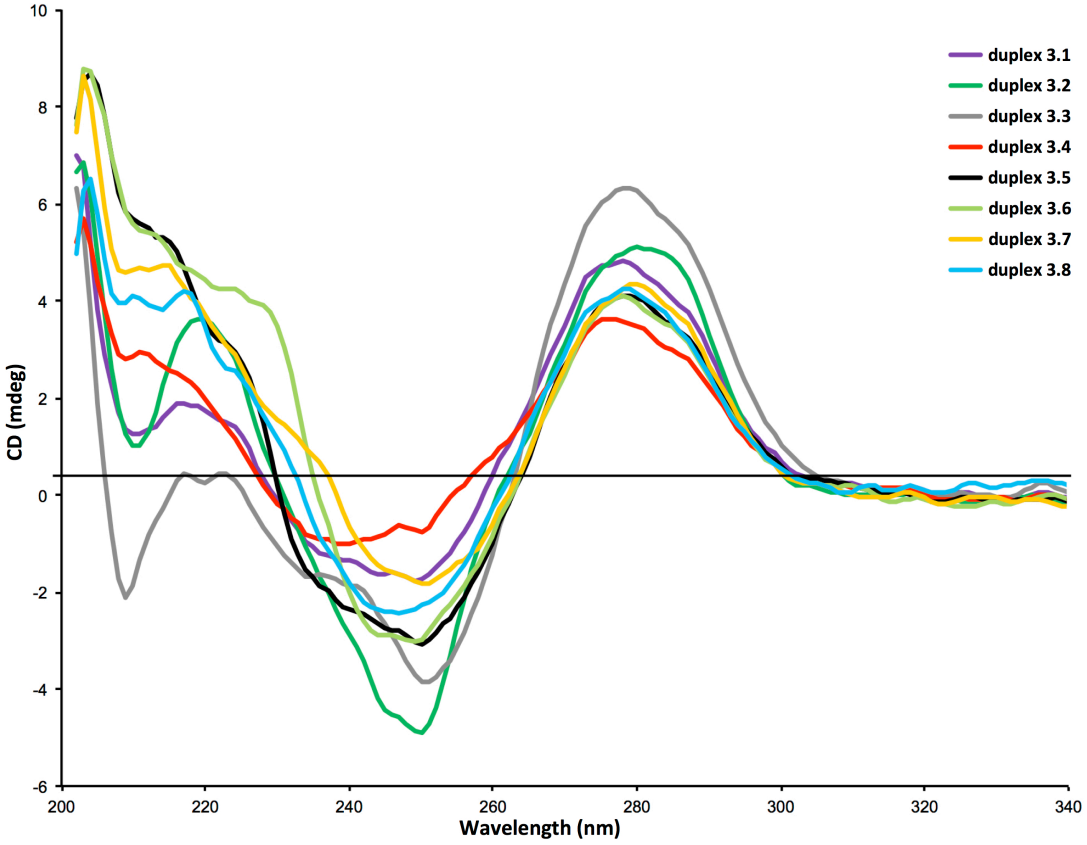


Figure 3.18 CD spectra of **duplexes 3.1-3.8**. All spectra were recorded at a duplex concentration of 1.5 μM (pH 7, 100 mM NaCl, 10 mM NaH_2PO_4, 0.1 mM EDTA).

3.4. Discussion

Our results indicate a strong preference for alternating NDI-DAN-NDI stacking relative to NDI-NDI-NDI and especially DAN-DAN-DAN self-stacking within a DNA duplex. Comparing **duplex 3.5** with **duplex 3.6**, an impressive 10°C increase in melting

temperature was observed upon replacement of the central DAN with NDI to create the alternating DAN-NDI-DAN arrangement. Consistent with this trend, the melting temperature of **duplex 3.7** compared with **duplex 3.8** demonstrates that there was a 4°C increase upon changing the central NDI to a DAN resulting in the alternating NDI-DAN-NDI arrangement. Also consistent with our current understanding is that the latter difference is less significant than the former because NDI is known to self-stack with relatively high stability, albeit in an offset stacking interaction, while DAN self-stacking is not known to be favorable.[55] It is therefore demonstrated that the known electrostatic complementarity between these units is stabilizing in the context of DNA duplexes. It should also be noted that each of the DAN and NDI unit is paired across from a "spacer" methyl unit, which is void of aromatic character and hydrogen bonding ability therefore limiting the driving forces normally considered in DNA duplex formation.

The comparable melting temperatures of **duplex 3.8** and **duplex 3.2** are surprising considering that three aromatic units are removed from the three base-pair region which decreases the desolvation-driven pairing of aromatic units and removes the stabilizing hydrogen bonding interactions seen between base pairs. Additionally, it is seen that **duplex 3.8** had a 5°C higher melting temperature than **duplex 3.6**, consistent with the observation that NDI is a known DNA intercalator,[126] so the two NDI-nucleobase interactions predicted to occur at either end of the NDI-DAN-NDI segment within **duplex 3.8** should be energetically favorable compared with the two DAN-nucleobase interactions predicted to occur at either end of the DAN-NDI-DAN segment within **duplex 3.6**.

Although no direct spectroscopic evidence for NDI and DAN stacking in any of the duplexes was obtained, the predictable differences in stabilities, consistent with the

known preference for alternating NDI-DAN stacking, provides supporting evidence that especially in **duplex 3.8** the NDI and DAN units are stacked in, or close to, the preferred face-centered geometry. Based on CD analysis, we do know that the modified bases do not cause any significant deviations from B-form structure of the entire duplex. Overall, these considerations indicate the design of non-natural base surrogates **2.1** and **2.2** was able to facilitate the desired NDI-DAN stacking within the context of a B-form DNA helix.

3.5. Conclusion

In summary, two new DNA base surrogate building blocks based on DAN and NDI have been incorporated into DNA oligonucleotide strands via standard solid phase oligonucleotide synthesis. In the most favorable arrangement, an NDI-DAN-NDI sequence in the middle of a DNA duplex was found to be approximately as stabilizing as three A-T base pairs. The results demonstrate that electrostatic complementarity between DAN and NDI can be used to drive the specificity and stability of DNA duplexes. Future work will explore how NDI and DAN units behave in alternate patterns within a DNA duplex, including their effect at the terminal position of oligonucleotide duplexes, as well as potential oligo-DAN and oligo-NDI duplex formation.

3.6. Experimental Methods

3.6.1. "Spacer" Phosphoramidite Monomer Synthesis

(*S*)-1-(bis(4-methoxyphenyl)(phenyl)methoxy)propan-2-ol (3.1). Compound 3.1 was synthesized according to a protocol previously reported in the literature.[129] ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 7.44 (d, J = 7.1 Hz, 2H), 7.31 (dd, J = 15.4, 8.4 Hz, 6H), 7.25 – 7.16 (m, 1H), 6.84 (d, J = 9.0 Hz, 4H), 4.03 – 3.93 (m, 1H), 3.80 (s, 6H), 3.13

(dd, J = 9.2, 3.4 Hz, 1H), 3.00 (dd, J = 9.2, 7.9 Hz, 1H), 2.42 (s, 1H), 1.11 (d, J = 6.4 Hz, 3H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 158.4, 144.8, 136.0, 130.0, 129.1, 128.1, 127.8, 126.8, 113.1, 86.0, 68.8, 67.1, 55.2, 18.9. HRMS-ESI predicted for $C_{24}H_{26}O_4Na^+$ $[M+Na]^+$: calcd 401.1723 *m/z*, found 401.1727 *m/z*.

(*S*)-1-(bis(4-methoxyphenyl)(phenyl)methoxy)propan-2-yl (2-cyanoethyl) diisopropylphosphoramidite (3.2). In a dry 50 mL round bottom flask, **3.1** (0.2608 g, 0.6891 mmol) was dissolved in CH_2Cl_2 (12.3 mL) and *N*,*N*-diisopropylethylamine (0.77 mL, 4.4102 mmol) was added. The reaction was stirred and purged with argon and 2-cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.31 mL, 1.3897 mmol) was added dropwise. The reaction was stirred for 3.5 hours at room temperature. The reaction mixture was poured into saturated aq. $NaHCO_3$, washed 3 times with CH_2Cl_2 and dried over Na_2SO_4. The CH_2Cl_2 was removed in vacuo and the product was purified by silica gel column chromatography (1:4 Hexanes/Ethyl acetate with 0.1% triethylamine) to yield a mixture of diastereomers as a faint yellow oil (0.3146 g, 0.5437 mmol, 79% yield). ^{1}H NMR (400 MHz, $CDCl_3$, ppm) δ 7.48 (d, J = 8.3 Hz, 2H), 7.36 (dd, J = 8.9, 3.7 Hz, 4H), 7.28 (dd, J = 14.6, 7.4 Hz, 2H), 7.20 (dd, J = 14.7, 8.6 Hz, 1H), 6.83 (t, J = 8.5 Hz, 4H), 4.21 – 4.04 (m, 1H), 3.91 – 3.82 (m, 1H), 3.79 (d, J = 3.5 Hz, 6H), 3.77 – 3.72 (m, 1H), 3.70 – 3.51 (m, 2H), 3.22 – 3.14 (m, 1H), 3.02 (dd, J = 9.4, 4.6 Hz, 1H), 2.91 (dd, J = 9.0, 5.7 Hz, 1H), 2.62 (t, J = 6.5 Hz, 1H), 2.56 – 2.41 (m, 1H), 1.28 (t, J = 5.3 Hz, 3H), 1.24 – 1.16 (m, 9H), 1.11 (d, J = 6.8 Hz, 3H). ^{13}C NMR (400 MHz, $CDCl_3$, ppm) δ 158.3, 145.0, 136.3, 136.2, 130.1, 130.0, 129.1, 128.2, 127.7, 127.6, 126.6, 126.5, 117.7, 113.1, 112.9, 85.8, 85.7, 70.0, 69.8, 68.0, 67.8, 67.8, 58.5, 58.3, 58.1, 57.9, 55.1, 43.0, 42.9, 24.7, 24.6, 24.5, 24.4, 20.4, 20.3, 20.2, 19.7, 19.6. ^{31}P NMR (400 MHz, $CDCl_3$, ppm) δ 147.7, 147.4.

HRMS-ESI predicted for $C_{33}H_{44}N_2O_5P^+$ $[M+H]^+$: calcd 579.2982 *m/z*, found 579.2977 *m/z*.

3.6.2. Oligonucleotide Synthesis

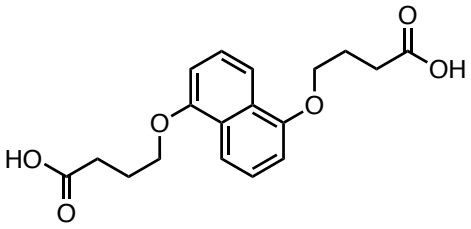
Unmodified oligonucleotides were synthesized on an automated nucleic acid synthesizer using a standard protocol for 2-cyanoethyl phosphoramidites (0.067 M) using a Glen UnySupport Expedite format column on a 1 μmol scale. Due to poor solubility in CH_3CN, all modified base phosphoramidites were diluted with 3:1 CH_2Cl_2:CH_3CN. The oligonucleotides were synthesized using a trityl-on synthesis and cleaved off of the resin with 1 mL concentrated aqueous ammonia at room temperature for 12-24 hours. The cleaved oligonucleotides were diluted with 1 mL NaCl solution (100 mg/mL) and then semi-purified by application to a Glen Pak purification column.

To avoid aminolysis, oligonucleotides containing NDI were synthesized utilizing UltraMild synthesis and deprotection methods from Glen Research. The UltraMild compatible phosphoramidites were used (Pac-dA-CE, Ac-dC-CE, and iPr-Pac-dG-CD Phosphoramidites). The Universal Support III expedite format column on a 1 μmol scale was used. The oligonucleotides were cleaved off of the resin using the UltraMild deprotection solution (0.05 M K_2CO_3 in methanol) for 12 hours then diluted with 1 mL of 0.1 M TEAA and neutralized with 6 μL glacial acetic acid and lyophilized. The resulting dried crude trityl-on oligonucleotides were dissolved in 2 mL 0.1 M TEAA and then applied directly on a Glen Pak purification column. Standard protocol for trityl-on oligonucleotides was used for the Glen Pak purification columns to afford the semi-purified detritylated oligonucleotides. All semi-purified oligonucleotides were lastly purified by reverse phase HPLC using a C18 peptide semipreparatory reverse-phase column with 0.1 M aqueous TEAA (pH 7) and CH_3CN as the eluent. Oligonucleotides

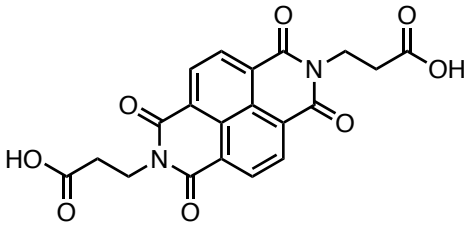
were all characterized by HRMS-ESI (negative mode, CH_3CN/aqueous ammonium carbonate).

3.6.3. Determining Extinction Coefficient of 2.1 and 2.2

The extinction coefficient of **2.1** and **2.2** were determined based on dilutions and UV measurements of water-soluble DAN and NDI analogues (Figure 3.19). The DAN and NDI analogues were diluted with a phosphate buffer (pH 7, 100 mM NaCl, 10 mM NaH_2PO_4, 0.1 mM EDTA) and UV absorbance measurements were taken at 260 nm. The resulting extinction coefficients (2504 and 1955 $M^{-1} \bullet cm^{-1}$) were used as an estimate for the corresponding phosphoramidites **2.1** and **2.2**, respectively. Because the content of the non-natural DNA base analogues is small compared to the amount of naturally occurring DNA bases, this approximation is unlikely to cause any major discrepancies in DNA concentration.



4,4'-(naphthalene-1,5-diylbis(oxy))dibutyric acid



3,3'-(1,3,6,8-tetraoxo-1,3,6,8-tetrahydrobenzo[*lmn*][3,8]phenanthroline-2,7-diyl)dipropionic acid

Figure 3.19 Structures of (left) DAN and (right) NDI analogues used to determine extinction coefficients for **2.1** and **2.2**.

3.6.4. Thermal Denaturing Studies

Samples of each oligonucleotide strand were prepared at a concentration of 3 μM in phosphate buffer (pH 7, 100 mM NaCl, 10 mM NaH_2PO_4, 0.1 mM EDTA). The oligonucleotide concentrations were quantified by measuring absorbance at 260 nm. The

corresponding molar extinction coefficients were calculated by summing the individual extinction coefficients for all the bases in the sequence. The molar extinction coefficients for dA, dG, dT, dC, **2.1**, **2.2**, and **3.1** were taken as 15400, 11500, 8700, 7400, 2504, 1955 and 0, $M^{-1} \bullet cm^{-1}$ at 260 nm, respectively. Because the content of the non-natural DNA base analogues is small compared to the amount of naturally occurring DNA bases, this approximation is unlikely to cause any major discrepancies in DNA concentration. The melting studies were measured in a Teflon-stoppered 1-cm-path-length quartz cell on a UV-vis spectrophotometer equipped with a thermoprogrammer. Each melting temperature run involved combining 0.5 mL of complementary strand samples to obtain 1.5 μM of duplex in solution. The samples were initially heated to 85 °C for 5 min, and then cooled from 85°C to 70°C at a rate of 1°C/min and then from 70 °C to 5 °C at a rate of 0.5°C/min. The absorbance at 260 nm was monitored. Two runs of these experiments were carried out per sample and averaged.

3.6.5. Circular Dichroism Spectroscopy

CD spectra were performed on a circular spectropolarimeter, and all spectra were measured in a 1-cm-path-length quartz cell. Samples were prepared from UV-melting temperature studies (1.5 μM in duplex concentration) in a phosphate buffer (pH 7, 100 mM NaCl, 10 mM NaH_2PO_4, 0.1 mM EDTA). Samples were initially heated to 85 °C for 5 minutes and allowed to cool to room temperature for 45 minutes before collecting CD data. Experiments were carried out at 25 °C.

This chapter adapted with permission from Ikkanda, B.A., Samuel, S.A., and Iverson, B.L. (2014). NDI and DAN DNA: Nucleic Acid-Directed Assembly of NDI and

CHAPTER 4

Position Dependent Effects on Stability of NDI and DAN-Modified Oligonucleotide Duplexes

4.1. CHAPTER SUMMARY

4.1.1. Introduction

The successful incorporation of NDI and DAN phosphoramidites into an oligonucleotide of natural DNA base pairs and subsequent annealing and melting experiments was a step towards understanding the interaction of NDI and DAN within the context of nucleotidomimetic foldamers in aqueous environments. The NDI-DAN-NDI sequence afforded moderate stability to an oligonucleotide (approximately equivalent to three A-T base pairs). However, the observed temperature independence of the oligo-NDI and oligo-DAN heteroduplex system sparked the curiosity of a possible inherent property of heteroduplexes of NDI and DAN. To examine this further, a deeper understanding of DNA thermal denaturing was explored and new modified oligonucleotide constructs were designed. Even though the understanding of DNA melting kinetics and thermodynamics is continually deepening, a dominant model of DNA melting at high temperature of short oligonucleotides involves the fraying of strands at the oligonucleotide termini and subsequent cooperative "unzipping" to yield the random-coiled single strands. It is reasonable to predict that the maximal thermal stabilizing effect of a modification to an oligonucleotide would be at the ends of the oligonucleotide versus in the middle. Therefore, new oligonucleotides were designed to examine the positional dependence of DAN and NDI modifications to an oligonucleotide on its thermal stability.

4.1.2. Scientific Question

The aim of this chapter is to answer the question: *How will duplex stability be affected by placing an NDI-DAN-NDI triplet sequence at the termini of an oligonucleotide versus at an internal position?* The results of such a study would give greater insight into the influence of NDI-DAN-NDI interactions on heteroduplexes in water, and how big a role position plays in determining that influence.

4.1.3. Approach

A 9 base pair sequence was used as a conserved sequence to compare with two modified duplex designs where the NDI-DAN-NDI triplet sequence was placed either at the ends of the duplex or at an internal position within the duplex (generating a 15 base pair modified duplex). These constructs were compared with control oligonucleotides that consisted of analogous 15-mer sequence with three G-C base pairs or three A-T base pairs inserted in the corresponding positions. The stability of the duplexes were examined by UV thermal melting experiments and the helical nature of the duplexes was examined by circular dichroism spectrophotometry.

4.1.4. Results

The NDI-DAN-NDI modified duplexes demonstrated a significant position dependent effect on the DNA duplex stability, raising the melting temperature by 17ºC when comparing the internal modification to the terminal modified duplexes. For comparison, changing the position of the A-T or G-C sequences from internal to terminal positions had only a marginal effect on duplex thermal stability (±1 ºC). Additionally, the duplex containing the NDI-DAN-NDI sequence on terminal positions demonstrated the highest melting temperature of all the examined duplexes. All oligonucleotides and modified oligonucleotides generally displayed characteristic B-form DNA duplex CD

traces; however, the NDI-DAN-NDI internally modified duplex demonstrated a unique positive absorbance at 220 nm when compared to control duplexes indicating an apparent disruption to the natural B-form duplex structure.

4.2. BACKGROUND

4.2.1. DNA Duplex Stability

Not only is the DNA double helix a marvel of molecular architecture, but the various interactions that drive the double helical architecture are such that DNA can serve as a biologically useful molecular assembly. If the intermolecular interactions were too weak within DNA, the double helix might not ever fully form; however, if the intermolecular interactions were too strong, then the function of DNA would be seriously crippled. DNA must separate (denature) so that it can be replicated or transcribed into RNA, which is then translated into proteins. The denaturing of double-stranded DNA into its component single strands can be achieved in several ways: changes in pH, raising the temperature, the addition of denaturants or by means of enzymatic replication or transcription.

The official melting temperature of a DNA sequence is defined as the temperature at which the probability of observing hybridized double-stranded and separated single-stranded DNA is the same. The melting temperature can therefore be used as a relative measure of duplex stability provided conditions are the same for each measurement. The melting temperature of DNA is primarily a function of base pair composition and sequence, but also of experimental conditions such as pH, duplex concentration, and salt concentration. Due to the significant increase in absorption of UV light (at 260 nm) from the double helical structure to the denatured random coils of DNA, monitoring the

absorbance at 260 nm is a method to measure the melting temperature of DNA (Figure 4.1).[130]

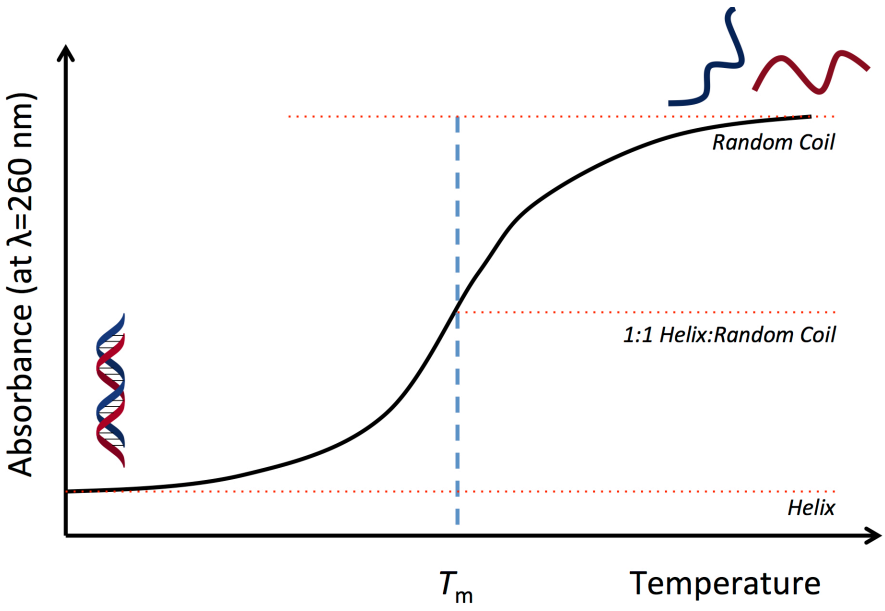


Figure 4.1 Thermal denaturing graph showing the helix to random coil transition and resulting melting temperature.

The primary interactions that maintain the double helical architecture of DNA are often described as twofold: hydrogen bonding and base stacking. By first inspection, these two components are reasonable and intuitive by observing the structure of DNA. The horizontal, in-plane contacts between nucleobases are described by Watson-Crick hydrogen bonding patterns while the vertical contacts between adjacent base pairs involve interactions associated with a geometry analogous to a deck of cards stacked with a twist. It is within these two points of contact that one observes the most direct inter-strand contacts in the DNA duplex. However, while the relative amounts of varying contributions of intermolecular forces stabilizing the double helix is still currently being

investigated, much research has advanced our understanding of the complexity of duplex stability in water, as briefly discussed below.

4.2.1.1. Hydrogen Bonding

The classical Watson-Crick hydrogen-bonding patterns that define the specificity for A-T and G-C base pairs have traditionally been assumed to be responsible for the stability of the DNA double helix. And indeed, when observing the melting temperature of double stranded DNA, the dominant contributing factor to the melting temperature is the G-C base pair content of the DNA sequence, suggesting that hydrogen bonding (3 hydrogen bonds for G-C base pairs compared to 2 hydrogen bonds for A-T base pairs) is the dominant contributor to duplex stability (Figure 4.2).

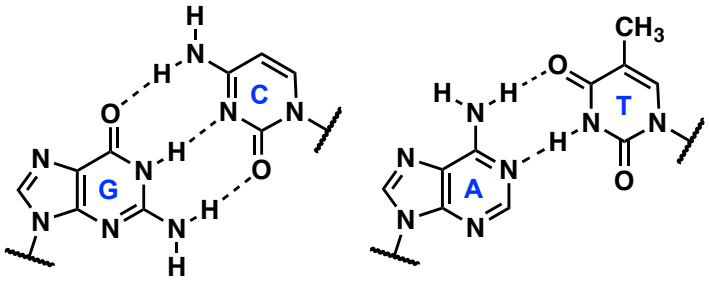


Figure 4.2 Watson-Crick hydrogen-bonding patterns.

However, solely considering the hydrogen-bonding capacity of the base pairs would be at odds with the observation of the potential numerous other hydrogen-bonding contacts that could be formed with the surrounding water molecules. When double-stranded DNA is denatured into single-stranded form, the number of hydrogen bonding interactions are essentially doubled as each base is not confined to hydrogen bonding with its pair but can freely hydrogen bond with water in solution.

Furthermore, recent work by Kool *et al*. has demonstrated that by utilizing an isosteric thymine analogue with significantly reduced hydrogen-bonding capability,

difluorotoluene (F), the DNA duplex is destabilized (Figure 4.3).[111] Their analysis suggests that whatever small favorable interaction there may be with F-A is overwhelmed by a large desolvation cost of the more polar adenine to bury itself within the double helical core of DNA when paired with F. Therefore, additional factors must be taken into consideration.

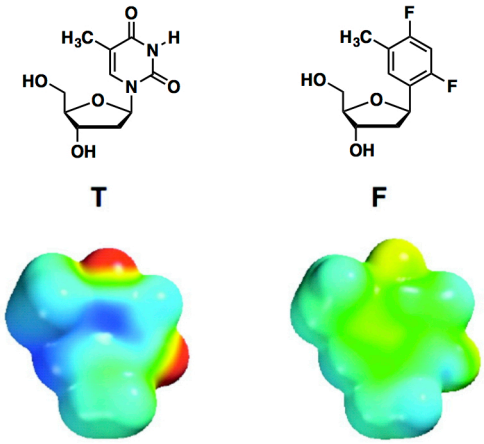


Figure 4.3 (*left*) structure of natural DNA nucleoside thymidine (T) with electrostatic potential map and (*right*) structure of isosteric non-natural nucleoside difluorotoluene (F). Adapted with permission from Reference 111. Copyright 2012 American Chemical Society.

4.2.1.2. Base Stacking

The stacking geometry of the DNA bases provides a context for additional contributing forces in duplex stability (van der Waals, electrostatic forces, and solvation/desolvation, i.e. the hydrophobic effect, for example). While the influence of hydrogen bonding on DNA stability is mainly dictated by the composition (GC content) of a sequence of DNA, base stacking does not only take into consideration the composition of bases but also their sequence. The relative contributions of these particular forces are complex and not fully understood; however, some recent work in the

field is able to shed light into the different contributions of base-stacking with the greatest influence on duplex stability.

As described in Chapter 1, there does seem to be an energetically favorable dispersion component to aromatic units stacking on top of each other, which is also suggested to stabilize the DNA double helix. Computational work by Honig *et al.* evaluated the contributions of electrostatic interactions, van der Waals interactions, and solvation/desolvation on base stacking by using continuum solvation methods.[131-132] By evaluating the free energy of association of nucleic acid bases to form varying stacked complexes, Honig concluded that the electrostatic contribution is minimal or even repulsive in the stacked conformation while dispersion and hydrophobic effects dominate the stacking geometry. However, conformational arrangement of stacked species studied was taken to be of the particular form determined from parameters defined by the ideal B-form architecture (rise = 3.4 Å, twist = 36, and all other helical parameters = 0), which may neither constitute the ideal electrostatic conformation of stacked bases nor actual conformation of realistic stacked bases in solution. Furthermore, the electrostatic cost of base stacking was not due to favorable/unfavorable interactions between base stacks but rather the energetic penalty of losing favorable electrostatic base-solvent interactions when in the stacked conformation. Nevertheless, from their study what was clear was van der Waals interactions were evident by the dipole/induced dipole interactions calculated when bases were close-packed in the stacking geometry. Furthermore, these computer modeling calculations clearly showed that solvation/desolvation plays a significant role in base stacking.

Due to the nature of electronegative atoms contained in and appended to the heterocyclic nucleobases, it is also proposed that electrostatic complementarity may be a

driving force for base stacking geometries. Indeed, in various crystal and co-crystal structures of nucleobases and modified nucleobases, a stacked geometry is observed.[133] Furthermore, the repeated patterns of stacking conformations in the solid state, even with modifications of various substituents to the nucleobases, demonstrate a favorable electrostatic preference for the stacking conformation.[133] What is commonly observed is that substituents (amine nitrogen, carbonyl oxygen or halogens) of one nucleobase seem to preferentially be placed over the aromatic system of the adjacent base.

Additionally, a large hydrophobic driving force is operating as evidenced by the placement of the relatively non-polar aromatic bases stacked within the interior of the DNA double helix in aqueous environments. In order to study the effect of stacking (as opposed to hydrogen bonding) on the stability of a DNA duplex, Kool and coworkers utilized a series of "dangling end" oligonucleotides as model systems (Figure 4.4).[110] These self-complementary short oligonucleotides place an unpaired nucleobase or aromatic unit at both ends. The dangling end duplexes are a model system that enables one to study the effect of introducing isolated stacking interactions on duplex stability.[134] Observing the thermal denaturing curves of various aromatic "dangling ends" in aqueous buffer, the larger non-polar aromatic molecules such as pyrene tended to stabilize the DNA duplex the most. Pyrene is also the most affected by the addition of a more non-polar solvent (ethanol) indicating that for large, non-polar aromatic molecules, solvation/desolvation is the major factor contributing to duplex stability.

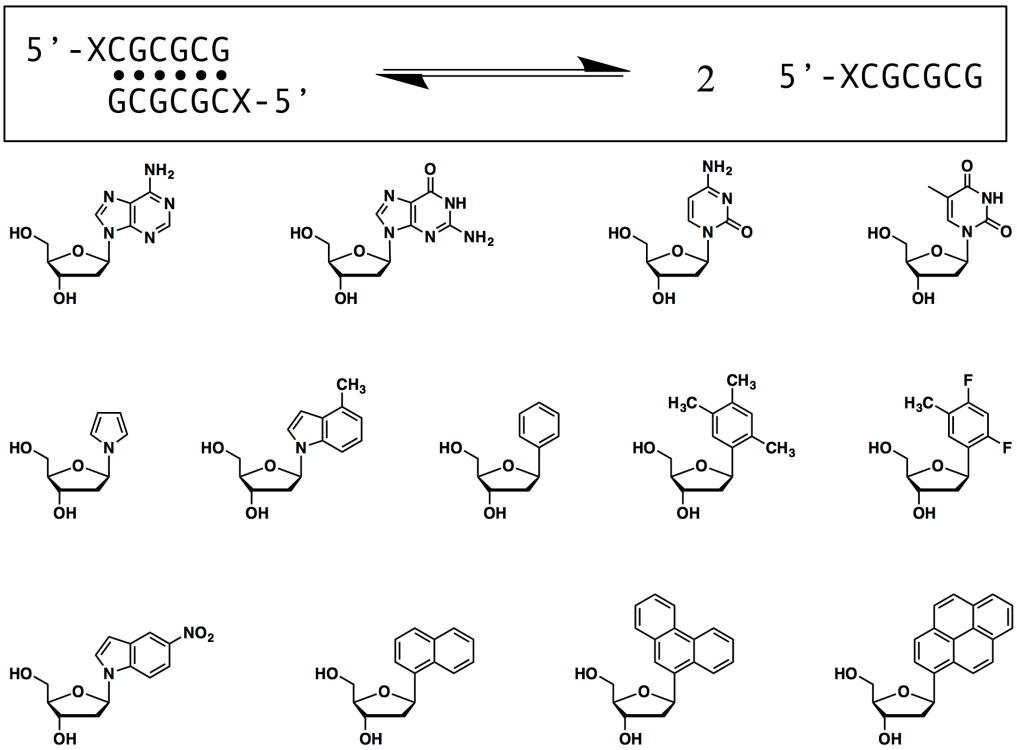


Figure 4.4 Design of "dangling end" experiment and the various natural and unnatural nucleosides incorporated as the dangling end, "X".[110]

The emerging picture of the various non-covalent interactions that stabilize the DNA double helix is that these three forces (van der Waals, electrostatic interactions, and especially solvation/desolvation) all play a role in the stacking geometry and stabilization of the duplex structure. However, electrostatic and van der Waals interactions seem to play a more minimal role as a stacking driving force and is difficult to determine their effect separate from solvation/desolvation, which seems to be a major contributor to the duplex structure and stability.

4.2.2. Sequence Dependence on Thermal Denaturation of DNA

Experimentally and theoretically determined parameters have been developed and provide the tools for a useful model to predict the melting temperature of DNA deemed the "nearest-neighbor" model pioneered by Zimm[135] and by Tinoco and coworkers.[136-137] This model takes into account not only the composition of DNA base pairs within a sequence (often defined as GC content), but assigns the thermodynamic parameters associated with the interaction between nearest-neighbor base pair stacks.

The model theorizes that the total free energy of an oligonucleotide or polymeric DNA consists of the sum of the standard free energy changes for the 10 possible Watson-Crick nearest-neighbor interactions (base pair stacks) corrected with terms specific for having an A-T base pair or G-C base pair at the terminus [ΔG°(init w/ term G-C/A-T)] and corrected by the additional entropic costs of duplex formation of symmetrical self-complementary duplexes [ΔG°(sym)].

$$\Delta G_{37}{}^{\circ}(total) = \sum_i n_i \Delta G^{\circ}(i) + \Delta G^{\circ}(\text{init w/ term G-C}) + \Delta G^{\circ}(\text{init w/ term A-T}) + \Delta G^{\circ}(\text{sym})$$

Equation 4.1 Standard free energy of duplex formation.

The $\Delta G^{\circ}(i)$ term is the sum of the empirically determined standard free-energy changes for the 10 possible Watson-Crick nearest-neighbors. A number of thermodynamic studies using a variety of methods (UV thermal denaturing, isothermal titration calorimetry, differential scanning calorimetry, etc.) on a wide range of oligonucleotide sequences and lengths have each determined these parameters with seemingly wide variations. However, a unified view of DNA nearest-neighbor thermodynamic parameters has been put forth by SantaLucia who closely examined the

experimental conditions used in several different studies and compiles reasonable explanations for such variations (Table 4.1).[138] It was emphasized by SantaLucia *et al*.[139] and Gray[140] that an additional correction must be included that takes into consideration the attenuated stability afforded at the terminal position as opposed to an internal position, which although minor, affects the overall duplex thermal stability.

Sequence	ΔH^o (kcal • mol^{-1})	ΔS^o (cal • K^{-1} • mol^{-1})	ΔG^o_{37} (kcal • mol^{-1})
AA/TT	– 7.9	– 22.2	– 1.00
AT/TA	– 7.2	– 20.4	– 0.88
TA/AT	– 7.2	– 21.3	– 0.58
CA/GT	– 8.5	– 22.7	– 1.45
GT/CA	– 8.4	– 22.4	– 1.44
CT/GA	– 7.8	– 21.0	– 1.28
GA/CT	– 8.2	– 22.2	– 1.30
CG/GC	– 10.6	– 27.2	– 2.17
GC/CG	– 9.8	– 24.4	– 2.24
GG/CC	– 8.0	– 19.9	– 1.84
Init w/ term. G-C	0.1	– 2.8	0.98
Init w/ term. A-T	2.3	4.1	1.03
Symmetry Correction	0	– 1.4	0.4

Table 4.1 Nearest-neighbor thermodynamic parameters for Watson-Crick base pair formation in 1 M NaCl.[138]

From these thermodynamic parameters, coupled with corrections for duplex concentration and ion concentration, one can fairly accurately predict the melting temperature. An equation that relates the thermal denaturing melting point of DNA to thermodynamic parameters via the Van't Hoff analysis is utilized to estimate the melting point of oligonucleotides of a particular sequence. This analysis requires the assumption of a two-state model of DNA melting as well as temperature independent enthalpy and entropy terms (which may not always be accurate assumptions).

$$T_M = \Delta H°/(\Delta S° + R \ln C_T)$$

Equation 4.2 The determination of melting temperature (T_M) of DNA duplex. R is the ideal gas constant (1.987 cal • K^{-1} • mol^{-1}). C_T is the total oligonucleotide duplex concentration.

A word of caution: while this model provides a useful tool in predicting the thermal stability of DNA, there is still much to be learned about the factors defining duplex thermodynamics. Due to the cooperative nature of DNA duplex formation as well as the complex nature of competing and interacting forces and conditions affecting double helical conformation (not to mention entropy-enthalpy compensation observed in some examples of DNA duplex formation),[141] defining thermodynamic parameters is a complex and difficult pursuit. It has been observed that UV melting curves evaluated by the Van't Hoff analysis may produce thermodynamic parameters that are significantly different from calorimetric determined values.[142-143] Examining experimentally determined thermal denaturing UV curves is currently a useful means to present relative duplex thermal stability so they will be used in the present study of our NDI and DAN modified duplexes. No conclusions related to actual thermodynamic parameters will be made from such UV melting experiments.

4.2.3. DNA Melting Models

The observation of the DNA melting curve demonstrates obvious cooperativity in the hybridization and denaturing processes of the DNA duplex starting with a slow nucleation event followed by a quick zipping/unzipping of the duplex architecture. A model for duplex formation describes a process of three consecutive base pairs sequentially coming together providing adequate nucleation for the addition of the rest of the stacked bases to come together much like a zipper being closed.[133] With each

additional stacked base pair a lower energetic cost is demanded for subsequent base pair construction.

Similarly, for DNA thermal denaturation, a nucleation event occurs, primarily with the melting of A-T rich regions first followed by the rest of the duplex, especially in long oligonucleotides. However, for short oligonucleotides the fraying of the oligonucleotide termini is a more likely nucleation site for DNA melting. Molecular dynamics computational studies affirm such a model where the breaking of hydrogen bonds of the base pairs by the twisting motions of the bases at the end of the oligonucleotide is followed by unstacking of the base pairs. Subsequently, the frayed bases rupture the hydrogen bonding and stacking of the next base pair simultaneously.[144] Through molecular dynamics computational modeling, a model of DNA melting was put forth by Wong and Pettitt which involves three conformational states: searching, fraying, and peeling (Figure 4.5).[145] Searching involves the incremental interchange of Watson-Crick hydrogen bonds with adjacent base pairs resulting in a shift in relative sequence registry between strands. Fraying is the tendency for the termini of DNA duplexes to break Watson-Crick hydrogen bonding structure. Peeling is the motion of strand untwising while folding back on itself such that nucleobases form hydrogen bonds with the phosphate backbone of a strand driving the fast cooperative unzipping of oligonucleotide base pairs.

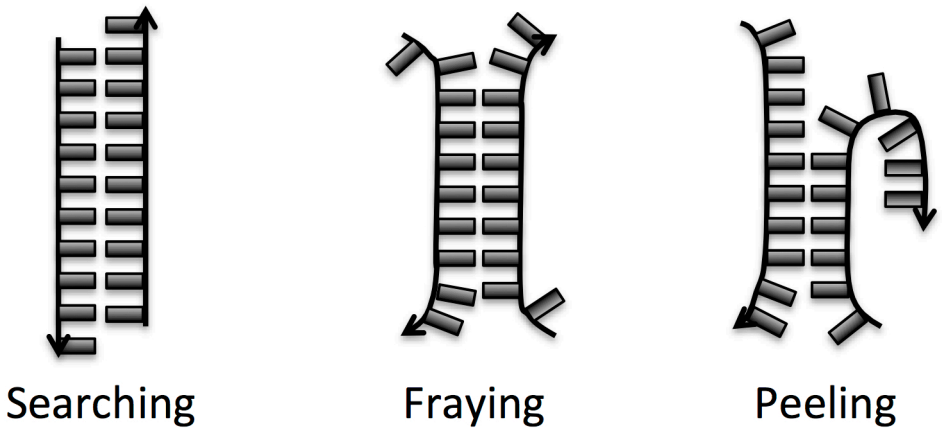


Figure 4.5 Schematic to show a proposed model of different conformational states of DNA melting (helical structure omitted for clarity)[145]

Previous work done by Greg Gabriel in our lab demonstrated that there is an apparent temperature independence to heteroduplex formation of the flexible peptide linked oligo-DAN and oligo-NDI. To further explore this observation, one would expect that the DAN and NDI interaction would have a dramatic effect on duplex thermal stability as shown by the melting temperature of DNA. As seen in the previous chapter, the NDI-DAN-NDI sequence, the most stable of the modified duplexes, only afforded moderate duplex thermal stability, approximately equivalent to three A-T base pairs. However, considering the sequence effects of base stacking on thermal stability as well as melting dynamics of oligonucleotides, the NDI-DAN-NDI thermal stability is hypothesized to have greatest impact when placed at the terminal positions of an oligonucleotide as opposed to an internal position.

4.3. RESULTS

4.3.1. Design

DAN (**2.1**), NDI (**2.2**), and spacer (**3.2**) phosphoramidites were used to append the DAN and NDI units to the strands of oligonucleotides using automated SPOS (Figure 4.6). An oligonucleotide of 9 base pairs, **duplex 4.1**, was chosen to be the conserved sequence of base pairs for this study (Figure 4.7). Position dependence was explored by placing the favorable NDI-DAN-NDI interaction either at the termini of the 9-mer DNA duplex, **duplex 4.7**, or three base pairs away from both ends of the duplex, **duplex 4.6**. Controls of analogous A-T (**duplex 2** and **duplex 3**) and G-C (**duplex 4** and **duplex 5**) base pair duplexes would also be examined to compare with natural DNA base pairing stability.

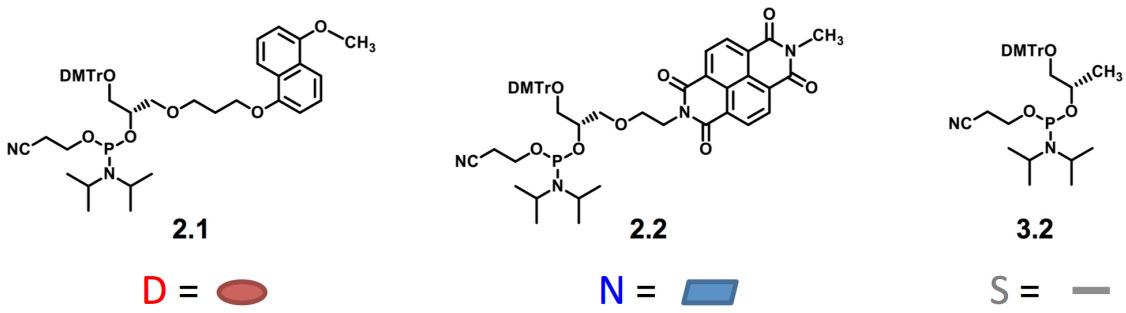


Figure 4.6 DAN, NDI, and spacer phosphoramidites used for the synthesis of modified oligonucleotides.

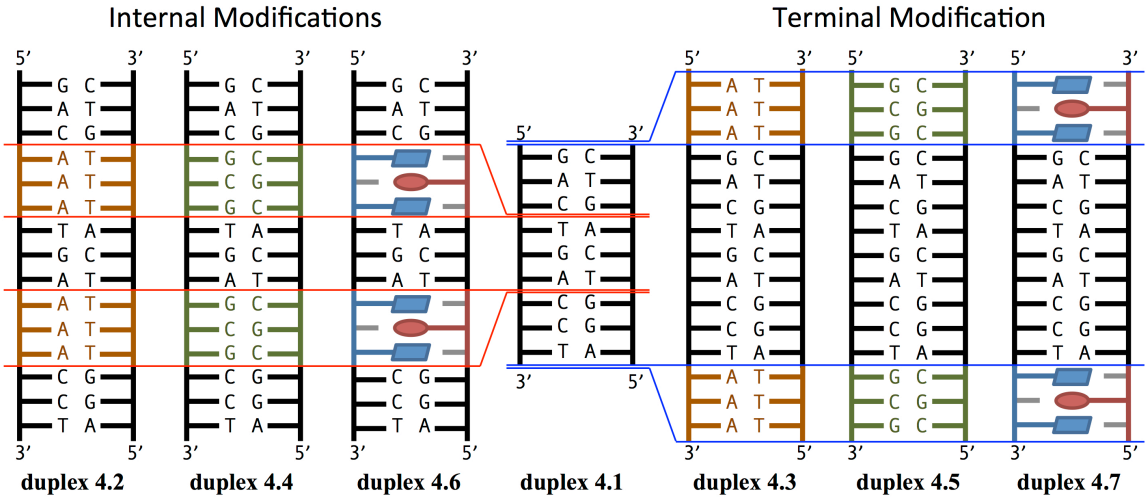


Figure 4.7 Duplex designs and control duplexes.

4.3.2. Synthesis

The phosphoramidite monomers for the DAN, NDI and spacer units were synthesized according to the improved synthetic routes described in Chapter 2. The NDI and DAN modified duplexes (**duplex 4.6** and **duplex 4.7**) were synthesized according to the protocol outlined in Chapter 3. The control duplexes (**duplex 4.1-4.5**) were purchased from IDT in the desalted form and further purified by reverse phase HPLC. The successful synthesis of the NDI and DAN modified oligonucleotides were confirmed by HRMS-ESI (Table 4.2)

Sequence	Ion	Mass (calcd)	Mass (found)
5′-NSN GAC TGA CCT NSN-3′	$[M-4H]^{-4}$	1202.4379	1202.4375
5′-SDS AGG TCA GTC SDS-3′	$[M-5H]^{-5}$	804.1406	804.1423
5′-GAC NSN TGA NSN CCT-3′	$[M-5H]^{-5}$	961.7489	961.7488
5′-AGG SDS TCA SDS GTC-3′	$[M-3H]^{-3}$	1340.9059	1340.9113

Table 4.2 HRMS-ESI of NDI and DAN modified oligonucleotides.

4.3.3. UV Thermal Denaturing Studies

Thermal denaturing studies were conducted to study the position dependent effect of the NDI-DAN-NDI modified duplexes (Table 4.3). All modifications increased the melting temperature beyond the control **duplex 4.1** by at least 15ºC. The duplexes with the internal and terminal sets of A-T base pairs (**duplex 4.2** and **duplex 4.3**) demonstrated the lowest of the melting temperatures while the duplexes with internally or terminally added G-C base pairs (**duplex 4.4** and **duplex 4.5**, respectively) had significantly higher melting temperatures by 14ºC and 16ºC, respectively. The position dependence of the all-natural DNA base-pairs was negligible, only differing by 1ºC when comparing internal (**duplex 4.2** and **duplex 4.4**) to terminal (**duplex 4.3** and **duplex 4.5**) substitution, respectively.

Duplex	Sequence	T_m (ºC)	
4.1	5′-GACTGACCT-3′ 3′-CTGACTGGA-5′	35	
4.2	5′-GAC **AAA** TGA **AAA** CCT-3′ 3′-CTG **TTT** ACT **TTT** GGA-5′	51	ΔT_m = -1 ºC
4.3	5′- **AAA** GACTGACCT **AAA** -3′ 3′- **TTT** CTGACTGGA **TTT** -5′	50	
4.4	5′-GAC **GCG** TGA **GCG** CCT-3′ 3′-CTG **CGC** ACT **CGC** GGA-5′	65	ΔT_m = +1 ºC
4.5	5′- **GCG** GACTGACCT **GCG** -3′ 3′- **CGC** CTGACTGGA **CGC** -5′	66	
4.6	5′-GAC **NSN** TGA **NSN** CCT-3′ 3′-CTG **SDS** ACT **SDS** GGA-5′	56	ΔT_m = +17 ºC
4.7	5′- **NSN** GACTGACCT **NSN** -3′ 3′- **SDS** CTGACTGGA **SDS** -5′	73	

Table 4.3 DNA melting temperatures. Duplex concentration of 1.5 μM (pH 7, 150 NaCl, 10 mM NaH_2PO_4, 10 mM EDTA)

Surprisingly, **duplex 4.7** had the highest melting temperature in the study, substantially higher than the G-C duplexes (**duplex 4.4** and **duplex 4.5**) by 8ºC and 7ºC, respectively. Furthermore, the position dependent effect on thermal stability was significant, an extraordinary 17ºC increase in melting temperature when comparing the internal positions (**duplex 4.6**) to the terminal positions (**duplex 4.7**). Interestingly, the internally modified NDI-DAN-NDI duplex (**duplex 4.6**) was not as stabilizing as the corresponding G-C control duplex, **duplex 4.4**, by 9ºC but was more stabilizing than the A-T control duplex, **duplex 4.2**, by 5ºC.

4.3.4. Circular Dichroism Spectroscopy

CD spectroscopy was used to probe structures of the duplexes used in this study. All modified duplexes generally demonstrated the characteristic B-form DNA bisignate CD spectrum in the region of 230 – 300 nm (Figure 4.8). It is noted that the NDI-DAN-NDI terminally modified duplex, **duplex 4.7**, has a roughly similar CD absorbance as the A-T control duplex, **duplex 4.3**. However, the NDI-DAN-NDI internally modified duplex (**duplex 4.6**) demonstrates a unique positive absorbance at 220 nm, which may indicate some structural deviation from the known B-form architecture of natural DNA.

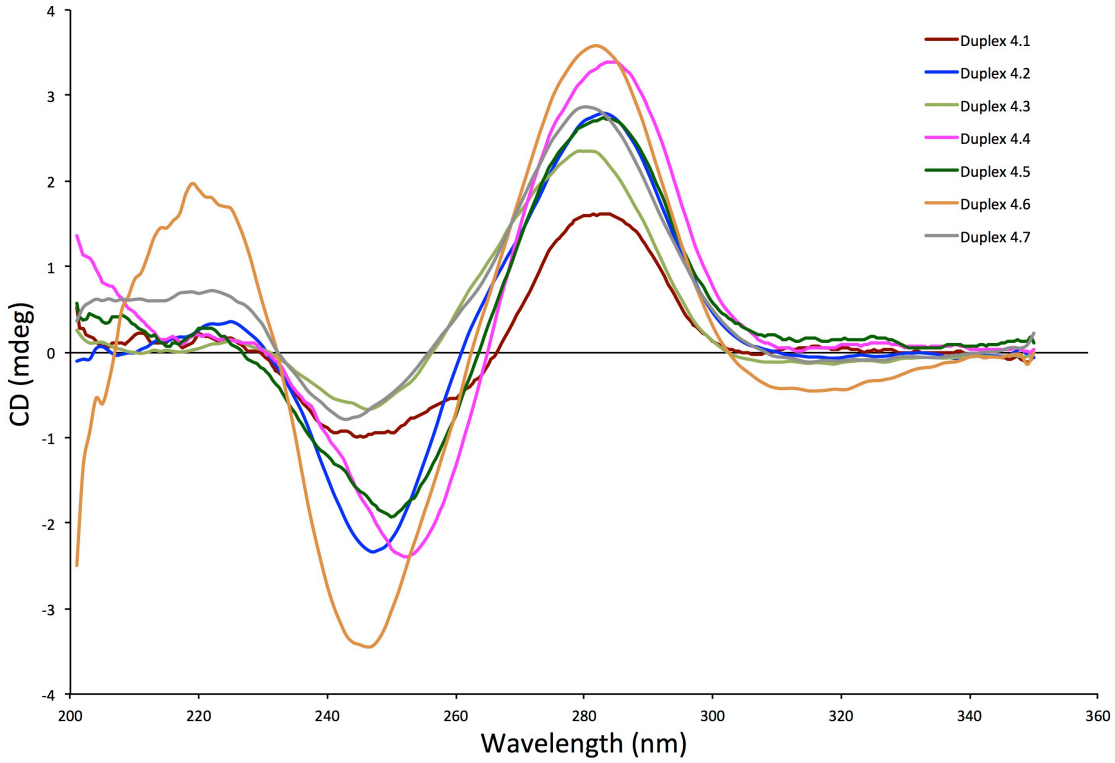


Figure 4.8 CD spectra of **duplexes 4.1-4.7**. All spectra were recorded at duplex concentrations of 1.5 μM (pH 7, 150 mM NaCl, 10 mM NaH_2PO_4, 10 mM EDTA).

Although CD spectroscopy is not a high-resolution structural technique, when comparing our modified duplexes with the control duplexes, the CD trace can be reasonably interpreted as indicating similar overall B-form structures for all but **duplex 4.6**. The unique features in the CD spectrum of **duplex 4.6** indicates the apparent disruption of the duplex structure to some extent when the NDI-DAN-NDI triplets are placed internally.

DISCUSSION

Taken together, our results indicate significant position dependence to the stability of DNA duplexes containing NDI-DAN-NDI triplets. When comparing the position dependence of the analogous three base pair sequence of G-Cs or A-Ts, the change in melting temperature is only by 1ºC. However, upon changing the internal NDI-DAN-NDI sequence to the terminal positions, a surprising 17ºC increase in melting temperature was observed. Note that if we consider **duplex 4.1** to be a reasonable control, a significant 38ºC increase in thermal stability is afforded (compared to 16° and 31° for A-T and G-C triplets, respectively). The position dependent effects of natural DNA base pairs as described by the nearest-neighbor model would suggest that there is only slight attenuation of the duplex stability added by the addition of a natural base-pair at the DNA termini. Therefore, *position alone* is likely not the key and there must be some other explanation(s) for the large stabilization observed with **duplex 4.7** compared to **duplex 4.6**.

The observation that **duplex 4.6** has a higher melting temperature than the analogous internal modifications of A-T base pairs, **duplex 4.2** (+5ºC), demonstrates that the NDI-DAN-NDI triplet sequence can be overall stabilizing to a DNA duplex (though not as stabilizing as three G-C base pairs). It should be noted that a single NDI-DAN-NDI triplet sequence within a DNA duplex had a 1ºC lower melting temperature than its analogous A-T base pair triplet sequence in our study in Chapter 3, while two NDI-DAN-NDI triplet sequences enables a DNA duplex to be *more* stable than it's analogous A-T base pair sequence. This may demonstrate that even greater stability is afforded when an oligonucleotide duplex is composed of more NDI-DAN-NDI interactions.

One could imagine at least two explanations, though they are not mutually exclusive. The first is that only the structural flexibility afforded by being at the termini allows adoption of the most favorable interacting conformation between the NDI and DAN aromatic units. The implication here is that when embedded in the middle of a B-form duplex, the NDI and DAN residues are held in positions that are not optimal for their most stable interacting geometry. A second explanation is that the presence of the NDI-DAN-NDI triplet in the middle of the duplex is destabilizing to adjacent base-pairs and/or the overall B-form duplex. According to this explanation, stability afforded by the internal NDI-DAN-NDI triplet is compromised by concomitant structural disruption of the duplex structure. Note that the highly unique features of the **duplex 4.6** CD spectrum might be consistent with either explanation, especially the latter one.

CONCLUSION

The position dependence of the NDI-DAN-NDI interaction on DNA duplex stability was examined through the construction of two oligonucleotide strands containing the NDI-DAN-NDI sequence at the ends of the oligonucleotide versus at an internal position of the sequence. The modified duplex with the NDI-DAN-NDI sequence at the terminal positions revealed a dramatic and surprising 17°C increase in melting temperature when compared to the analogous duplex placing the NDI-DAN-NDI sequence at internal positions. CD spectra provided evidence for some unique structural features of the duplex containing the internal NDI-DAN-NDI triplet, but not the highly stabilized terminal construct. Future studies will be conducted to further probe the scope, limitations and explanation for the duplex stabilizing NDI-DAN-NDI sequence.

4.4. EXPERIMENTAL METHODS

4.4.1. Oligonucleotide Synthesis

DAN (**2.1**), NDI (**2.2**), and spacer (**3.2**) phosphoramidites were synthesized and incorporated into oligonucleotides according to the procedures outlined in Chapter 2 and Chapter 3. Control oligonucleotides (**duplexes 4.1-4.5**) were purchased from IDT in desalted form and purified by reverse phase HPLC using a C18 peptide semipreparatory reverse-phase column with 0.1 M aqueous TEAA (pH 7) and CH_3CN as the eluent.

4.4.2. UV Thermal Denaturing Studies

Individual oligonucleotide strands were prepared at a concentration of 3 μM in phosphate buffer (pH 7, 150 mM NaCl, 10 mM NaH_2PO_4, 10 mM EDTA). Oligonucleotide concentrations were quantified by measuring the absorbance at 260 nm. Molar extinction coefficients were used as described in Chapter 3. The melting studies were measured in a screw capped 1-cm-path-length semi-micro quartz cell on a Cary 100 UV-vis spectrophotometer equipped with a thermoprogrammer. Each melting temperature run involved combining 0.4 mL of each complementary strand sample to obtain 1.5 μM of duplex in solution. The samples were heated to 90ºC and then cooled from 90ºC to 15º at a rate of 0.5ºC/min. The absorbance at 260 nm was monitored. The UV melting curves were then fitted according to a sigmoidal function and the derivative of the function was taken. The maximum of the resulting derivative plot was taken to be the melting temperature. At least two runs of these experiments were carried out per sample and averaged.

4.4.3. Circular Dichroism Spectroscopy

CD spectra were performed on a Jasco 815 circular spectropolarimeter, and all spectra were measured using the same samples [1.5 μM strand concentration in phosphate buffer (pH 7, 150 mM NaCl, 10 mM NaH_2PO_4, 10 mM EDTA)] and quartz cuvettes as described in the thermal denaturing studies. Samples were initially heated to 90ºC for 5 minutes and allowed to cool to room temperature for 1 hour before collecting CD spectra. Experiments were carried out at 20ºC. Each spectrum is the average of three sequential scans of the same sample. The resulting spectra were then smoothed using a moving average function and base line corrected.

References

1. von Ballmoos, C.; Wiedenmann, A.; Dimroth, P., Essentials for ATP synthesis by F1F0 ATP synthases. *Annu. Rev. Biochem* **2009,** *78*, 649-72.
2. Gellman, S. H., Foldamers: a manifesto. *Acc. Chem. Res.* **1998,** *31* (4), 173-180.
3. Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S., A Field Guide to Foldamers. *Chem. Rev.* **2001,** *101* (12), 3893-4012.
4. Appella, D. H.; Christianson, L. A.; Klein, D. A.; Powell, D. R.; Huang, X. L.; Barchi, J. J.; Gellman, S. H., Residue-based control of helix shape in beta-peptide oligomers. *Nature* **1997,** *387* (6631), 381-384.
5. Dado, G. P.; Gellman, S. H., Intramolecular Hydrogen-Bonding in Derivatives of Beta-Alanine and Gamma-Amino Butyric-Acid - Model Studies for the Folding of Unnatural Polypeptide Backbones. *J. Am. Chem. Soc.* **1994,** *116* (3), 1054-1062.
6. Krauthauser, S.; Christianson, L. A.; Powell, D. R.; Gellman, S. H., Antiparallel sheet formation in beta-peptide foldamers: Effects of beta-amino acid substitution on conformational preference. *J. Am. Chem. Soc.* **1997,** *119* (48), 11719-11720.
7. Hintermann, T.; Seebach, D., Synthesis of a beta-hexapeptide from (R)-2-aminomethyl-alkanoic acids and structural investigations. *Synlett* **1997,** (5), 437-438.
8. Seebach, D.; Gademann, K.; Schreiber, J. V.; Matthews, J. L.; Hintermann, T.; Jaun, B.; Oberer, L.; Hommel, U.; Widmer, H., 'Mixed' beta-peptides: A unique helical secondary structure in solution. *Helv. Chim. Acta* **1997,** *80* (7), 2033-2038.
9. Seebach, D.; Matthews, J. L., beta-peptides: a surprise at every turn. *Chem. Commun.* **1997,** (21), 2015-2022.
10. Seebach, D.; Overhand, M.; Kuhnle, F. N. M.; Martinoni, B.; Oberer, L.; Hommel, U.; Widmer, H., beta-peptides: Synthesis by Arndt-Eistert homologation with concomitant peptide coupling. Structure determination by NMR and CD spectroscopy and by X-ray crystallography. Helical secondary structure of a beta-hexapeptide in solution and its stability towards pepsin. *Helv. Chim. Acta* **1996,** *79* (4), 913-941.
11. Fisher, B. F.; Guo, L.; Dolinar, B. S.; Guzei, I. A.; Gellman, S. H., Heterogeneous H-bonding in a foldamer helix. *J. Am. Chem. Soc.* **2015,** *137* (20), 6484-7.
12. Horne, W. S.; Gellman, S. H., Foldamers with heterogeneous backbones. *Acc. Chem. Res.* **2008,** *41* (10), 1399-408.

13. Horne, W. S.; Price, J. L.; Keck, J. L.; Gellman, S. H., Helix bundle quaternary structure from alpha/beta-peptide foldamers. *J. Am. Chem. Soc.* **2007,** *129* (14), 4178-80.

14. Checco, J. W.; Lee, E. F.; Evangelista, M.; Sleebs, N. J.; Rogers, K.; Pettikiriarachchi, A.; Kershaw, N. J.; Eddinger, G. A.; Belair, D. G.; Wilson, J. L.; Eller, C. H.; Raines, R. T.; Murphy, W. L.; Smith, B. J.; Gellman, S. H.; Fairlie, W. D., alpha/beta-Peptide Foldamers Targeting Intracellular Protein-Protein Interactions with Activity in Living Cells. *J. Am. Chem. Soc.* **2015,** *137* (35), 11365-75.

15. Bassani, D. M.; Lehn, J. M.; Baum, G.; Fenske, D., Designed self-generation of an extended helical structure from an achiral polyheterocyclic strand. *Angewandte Chemie-International Edition in English* **1997,** *36* (17), 1845-1847.

16. Gardinier, K. M.; Khoury, R. G.; Lehn, J. M., Enforced helicity: Efficient access to self-organized helical molecular strands by the imine route. *Chemistry-a European Journal* **2000,** *6* (22), 4124-4131.

17. Ohkita, M.; Lehn, J. M.; Baum, G.; Fenske, D., Helicity coding: Programmed molecular self-organization of achiral nonbiological strands into multiturn helical superstructures: Synthesis and characterization of alternating pyridine-pyrimidine oligomers. *Chemistry-a European Journal* **1999,** *5* (12), 3471-3481.

18. Berl, V.; Huc, I.; Khoury, R. G.; Krische, M. J.; Lehn, J. M., Interconversion of single and double helices formed from synthetic molecular strands. *Nature* **2000,** *407* (6805), 720-723.

19. Zhang, D.-W.; Zhao, X.; Hou, J.-L.; Li, Z.-T., Aromatic Amide Foldamers: Structures, Properties, and Functions. *Chem. Rev.* **2012,** *112* (10), 5271-5316.

20. Sebaoun, L.; Maurizot, V.; Granier, T.; Kauffmann, B.; Huc, I., Aromatic Oligoamide beta-Sheet Foldamers. *J. Am. Chem. Soc.* **2014,** *136* (5), 2168-2174.

21. Stone, M. T.; Moore, J. S., A water-soluble m-phenylene ethynylene foldamer. *Org. Lett.* **2004,** *6* (4), 469-72.

22. Goto, K.; Moore, J. S., Sequence-specific binding of m-phenylene ethynylene foldamers to a piperazinium dihydrochloride salt. *Org. Lett.* **2005,** *7* (9), 1683-6.

23. Chongsiriwatana, N. P.; Patch, J. A.; Czyzewski, A. M.; Dohm, M. T.; Ivankin, A.; Gidalevitz, D.; Zuckermann, R. N.; Barron, A. E., Peptoids that mimic the structure, function, and mechanism of helical antimicrobial peptides. *Proceedings of the National Academy of Sciences of the United States of America* **2008,** *105* (8), 2794-2799.

24. Jun, J. M. V.; Altoe, M. V. P.; Aloni, S.; Zuckermann, R. N., Peptoid nanosheets as soluble, two-dimensional templates for calcium carbonate mineralization. *Chem. Commun.* **2015,** *51* (50), 10218-10221.

25. Malinovskii, V. L.; Wenger, D.; Haner, R., Nucleic acid-guided assembly of aromatic chromophores. *Chem. Soc. Rev.* **2010,** *39* (2), 410-422.

26. Zeng, H.; Ickes, H.; Flowers, R. A., 2nd; Gong, B., Sequence specificity of hydrogen-bonded molecular duplexes. *J. Org. Chem.* **2001,** *66* (10), 3574-83.

27. Gong, H.; Krische, M. J., Duplex molecular strands based on the 3,6-diaminopyridazine hydrogen bonding motif: amplifying small-molecule self-assembly preferences through preorganization and iterative arrangement of binding residues. *J. Am. Chem. Soc.* **2005,** *127* (6), 1719-25.

28. Bisson, A. P.; Carver, F. J.; Hunter, C. A.; Waltho, J. P., Molecular Zippers. *J. Am. Chem. Soc.* **1994,** *116* (22), 10292-10293.

29. Berl, V.; Huc, I.; Khoury, R. G.; Lehn, J. M., Helical molecular programming: Supramolecular double helices by dimerization of helical oligopyridine-dicarboxamide strands. *Chemistry-a European Journal* **2001,** *7* (13), 2810-2820.

30. Smith, V. C. M.; Lehn, J. M., Helicate self-assembly from heterotopic ligand strands of specific binding site sequence. *Chem. Commun.* **1996,** (24), 2733-2734.

31. Stiller, R.; Lehn, J. M., Synthesis and properties of silver(I) and copper(I) helicates with imine-bridged oligobipyidine ligands. *Eur. J. Inorg. Chem.* **1998,** (7), 977-982.

32. Scott Lokey, R.; Iverson, B. L., Synthetic molecules that fold into a pleated secondary structure in solution. *Nature* **1995,** *375* (6529), 303-305.

33. Hunter, C. A.; Sanders, J. K., The nature of. pi.-. pi. interactions. *J. Am. Chem. Soc.* **1990,** *112* (14), 5525-5534.

34. Cozzi, F.; Cinquini, M.; Annunziata, R.; Dwyer, T.; Siegel, J. S., Polar/Pi Interactions between Stacked Aryls in 1,8-Diarylnaphthalenes. *J. Am. Chem. Soc.* **1992,** *114* (14), 5729-5733.

35. Cozzi, F.; Annunziata, R.; Benaglia, M.; Cinquini, M.; Raimondi, L.; Baldridge, K. K.; Siegel, J. S., Through-space interactions between face-to-face, center-to-edge oriented arenes: importance of polar-pi effects. *Org Biomol Chem* **2003,** *1* (1), 157-62.

36. Cozzi, F.; Cinquini, M.; Annuziata, R.; Siegel, J. S., Dominance of Polar/Pi over Charge-Transfer Effects in Stacked Phenyl Interactions. *J. Am. Chem. Soc.* **1993,** *115* (12), 5330-5331.

37. Cozzi, F.; Ponzini, F.; Annunziata, R.; Cinquini, M.; Siegel, J. S., Polar Interactions between Stacked π Systems in Fluorinated 1,8-Diarylnaphthalenes: Importance of Quadrupole Moments in Molecular Recognition. *Angewandte Chemie International Edition in English* **1995,** *34* (9), 1019-1020.

38. Cozzi, F.; Siegel, J. S., Interaction between Stacked Aryl Groups in 1,8-Diarylnaphthalenes - Dominance of Polar/Pi over Charge-Transfer Effects. *Pure Appl. Chem.* **1995,** *67* (5), 683-689.

39. Williams, J. H.; Cockcroft, J. K.; Fitch, A. N., Structure of the Lowest Temperature Phase of the Solid Benzene Hexafluorobenzene Adduct. *Angewandte Chemie-International Edition in English* **1992,** *31* (12), 1655-1657.

40. Patrick, C. R.; Prosser, G. S., Molecular Complex of Benzene and Hexafluorobenzene. *Nature* **1960,** *187* (4742), 1021-1021.

41. Rashkin, M. J.; Waters, M. L., Unexpected substituent effects in offset pi-pi stacked interactions in water. *J. Am. Chem. Soc.* **2002,** *124* (9), 1860-1.

42. Wheeler, S. E.; Houk, K. N., Substituent effects in the benzene dimer are due to direct interactions of the substituents with the unsubstituted benzene. *J. Am. Chem. Soc.* **2008,** *130* (33), 10854-5.

43. Wheeler, S. E.; McNeil, A. J.; Muller, P.; Swager, T. M.; Houk, K. N., Probing substituent effects in aryl-aryl interactions using stereoselective Diels-Alder cycloadditions. *J. Am. Chem. Soc.* **2010,** *132* (10), 3304-11.

44. Wheeler, S. E., Local nature of substituent effects in stacking interactions. *J. Am. Chem. Soc.* **2011,** *133* (26), 10262-74.

45. Wheeler, S. E.; Bloom, J. W., Toward a more complete understanding of noncovalent interactions involving aromatic rings. *J. Phys. Chem. A* **2014,** *118* (32), 6133-47.

46. Ringer, A. L.; Sinnokrot, M. O.; Lively, R. P.; Sherrill, C. D., The Effect of Multiple Substituents on Sandwich and T-Shaped π–π Interactions. *Chemistry – A European Journal* **2006,** *12* (14), 3821-3828.

47. Sinnokrot, M. O.; Sherrill, C. D., Unexpected Substituent Effects in Face-to-Face π-Stacking Interactions. *The Journal of Physical Chemistry A* **2003,** *107* (41), 8377-8379.

48. Arnstein, S. A.; Sherrill, C. D., Substituent effects in parallel-displaced [small pi]-[small pi] interactions. *PCCP* **2008,** *10* (19), 2646-2655.

49. Snyder, S. E.; Huang, B.-S.; Chu, Y. W.; Lin, H.-S.; Carey, J. R., The Effects of Substituents on the Geometry of π–π Interactions. *Chemistry – A European Journal* **2012,** *18* (40), 12663-12671.

50. Lee, E. C.; Kim, D.; Jurečka, P.; Tarakeshwar, P.; Hobza, P.; Kim, K. S., Understanding of Assembly Phenomena by Aromatic–Aromatic Interactions: Benzene Dimer and the Substituted Systems. *The Journal of Physical Chemistry A* **2007,** *111* (18), 3446-3457.

51. Grimme, S., Do Special Noncovalent π–π Stacking Interactions Really Exist? *Angew. Chem. Int. Ed.* **2008,** *47* (18), 3430-3434.

52. Alvey, P. M.; Reczek, J. J.; Lynch, V.; Iverson, B. L., A systematic study of thermochromic aromatic donor-acceptor materials. *J. Org. Chem.* **2010,** *75* (22), 7682-90.

53. Ikkanda, B. A.; Samuel, S. A.; Iverson, B. L., NDI and DAN DNA: nucleic acid-directed assembly of NDI and DAN. *J. Org. Chem.* **2014,** *79* (5), 2029-37.

54. Reczek, J. J.; Villazor, K. R.; Lynch, V.; Swager, T. M.; Iverson, B. L., Tunable columnar mesophases utilizing C-2 symmetric aromatic donor-acceptor complexes. *J. Am. Chem. Soc.* **2006,** *128* (24), 7995-8002.

55. Cubberley, M. S.; Iverson, B. L., (1)H NMR investigation of solvent effects in aromatic stacking interactions. *J. Am. Chem. Soc.* **2001,** *123* (31), 7560-3.

56. Smithrud, D. B.; Diederich, F., Strength of Molecular Complexation of Apolar Solutes in Water and in Organic-Solvents Is Predictable by Linear Free-Energy Relationships - a General-Model for Solvation Effects on Apolar Binding. *J. Am. Chem. Soc.* **1990,** *112* (1), 339-343.

57. Martinez, C. R.; Iverson, B. L., Rethinking the term "pi-stacking". *Chemical Science* **2012,** *3* (7), 2191-2201.

58. Alonso, M.; Woller, T.; Martin-Martinez, F. J.; Contreras-Garcia, J.; Geerlings, P.; De Proft, F., Understanding the fundamental role of pi/pi, sigma/sigma, and sigma/pi dispersion interactions in shaping carbon-based materials. *Chemistry* **2014,** *20* (17), 4931-41.

59. Gabriel, G. J.; Sorey, S.; Iverson, B. L., Altering the folding patterns of naphthyl trimers. *J. Am. Chem. Soc.* **2005,** *127* (8), 2637-40.

60. Nguyen, J. Q.; Iverson, B. L., An Amphiphilic Folding Molecule That Undergoes an Irreversible Conformational Change. *J. Am. Chem. Soc.* **1999,** *121* (11), 2639-2640.

61. O'Shea, E. K.; Klemm, J. D.; Kim, P. S.; Alber, T., X-ray Structure of the GCN4 Leucine Zipper, a Two-Stranded, Parallel Coiled Coil. *Science* **1991,** *254* (5031), 539-544.

62. Bradford, V. J.; Iverson, B. L., Amyloid-like behavior in abiotic, amphiphilic foldamers. *J. Am. Chem. Soc.* **2008,** *130* (4), 1517-24.

63. Shao, H.; Gao, M.; Kim, S. H.; Jaroniec, C. P.; Parquette, J. R., Aqueous Self-Assembly of L-Lysine-Based Amphiphiles into 1D n-Type Nanotubes. *Chemistry – A European Journal* **2011,** *17* (46), 12882-12885.

64. Shao, H.; Nguyen, T.; Romano, N. C.; Modarelli, D. A.; Parquette, J. R., Self-Assembly of 1-D n-Type Nanostructures Based on Naphthalene Diimide-Appended Dipeptides. *J. Am. Chem. Soc.* **2009,** *131* (45), 16374-16376.

65. Shao, H.; Parquette, J. R., A [small pi]-conjugated hydrogel based on an Fmoc-dipeptide naphthalene diimide semiconductor. *Chem. Commun.* **2010,** *46* (24), 4285-4287.

66. Talukdar, P.; Bollot, G.; Mareda, J.; Sakai, N.; Matile, S., Ligand-Gated Synthetic Ion Channels. *Chemistry – A European Journal* **2005,** *11* (22), 6525-6532.

67. Talukdar, P.; Bollot, G.; Mareda, J.; Sakai, N.; Matile, S., Synthetic Ion Channels with Rigid-Rod π-Stack Architecture that Open in Response to Charge-Transfer Complex Formation. *J. Am. Chem. Soc.* **2005,** *127* (18), 6528-6529.

68. Avinash, M.; Govindaraju, T., A bio-inspired design strategy: Organization of tryptophan-appended naphthalenediimide into well-defined architectures induced by molecular interactions. *Nanoscale* **2011,** *3* (6), 2536-2543.

69. Pandeeswar, M.; Avinash, M. B.; Govindaraju, T., Chiral Transcription and Retentive Helical Memory: Probing Peptide Auxiliaries Appended with Naphthalenediimides for Their One-Dimensional Molecular Organization. *Chemistry – A European Journal* **2012,** *18* (16), 4818-4822.

70. Kar, H.; Molla, M. R.; Ghosh, S., Two-component gelation and morphology-dependent conductivity of a naphthalene-diimide (NDI) π-system by orthogonal hydrogen bonding. *Chem. Commun.* **2013,** *49* (39), 4220-4222.

71. Peebles, C.; Piland, R.; Iverson, B. L., More than meets the eye: conformational switching of a stacked dialkoxynaphthalene-naphthalenetetracarboxylic diimide (DAN-NDI) foldamer to an NDI-NDI fibril aggregate. *Chemistry* **2013,** *19* (35), 11598-602.

72. Gabriel, G. J.; Iverson, B. L., Aromatic oligomers that form hetero duplexes in aqueous solution. *J. Am. Chem. Soc.* **2002,** *124* (51), 15174-5.

73. Zhou, Q. Z.; Jiang, X. K.; Shao, X. B.; Chen, G. J.; Jia, M. X.; Li, Z. T., First zipper-featured molecular duplexes driven by cooperative donor-acceptor interaction. *Org. Lett.* **2003,** *5* (11), 1955-1958.

74. Zhou, Q. Z.; Jia, M. X.; Shao, X. B.; Wu, L. Z.; Jiang, X. K.; Li, Z. T.; Chen, G. J., Self-assembly of a novel series of hetero-duplexes driven by donor-acceptor interaction. *Tetrahedron* **2005,** *61* (30), 7117-7124.

75. Holman, G. G.; Zewail-Foote, M.; Smith, A. R.; Johnson, K. A.; Iverson, B. L., A sequence-specific threading tetra-intercalator with an extremely slow dissociation rate constant. *Nature Chemistry* **2011,** *3* (11), 875-881.

76. Rhoden Smith, A.; Iverson, B. L., Threading polyintercalators with extremely slow dissociation rates and extended DNA binding sites. *J. Am. Chem. Soc.* **2013,** *135* (34), 12783-9.

77. Smith, A. R.; Ikkanda, B. A.; Holman, G. G.; Iverson, B. L., Subtle recognition of 14-base pair DNA sequences via threading polyintercalation. *Biochemistry* **2012,** *51* (22), 4445-52.

78. Johnson, A. T.; Schlegel, M. K.; Meggers, E.; Essen, L.-O.; Wiest, O., On the Structure and Dynamics of Duplex GNA. *J. Org. Chem.* **2011,** *76* (19), 7964-7974.

79. Meggers, E.; Zhang, L., Synthesis and Properties of the Simplified Nucleic Acid Glycol Nucleic Acid. *Acc. Chem. Res.* **2010,** *43* (8), 1092-1102.

80. Schlegel, M. K.; Essen, L.-O.; Meggers, E., Duplex Structure of a Minimal Nucleic Acid. *J. Am. Chem. Soc.* **2008,** *130* (26), 8158-8159.

81. Schlegel, M. K.; Peritz, A. E.; Kittigowittana, K.; Zhang, L.; Meggers, E., Duplex formation of the simplified nucleic acid GNA. *ChemBioChem* **2007,** *8* (8), 927-932.

82. Zhang, L. L.; Peritz, A.; Meggers, E., A simple glycol nucleic acid. *J. Am. Chem. Soc.* **2005,** *127* (12), 4174-4175.

83. Zhang, L.; Peritz, A. E.; Carroll, P. J.; Meggers, E., Synthesis of glycol nucleic acids. *Synthesis* **2006,** (4), 645-653.

84. Reczek, J. J.; Villazor, K. R.; Lynch, V.; Swager, T. M.; Iverson, B. L., Tunable Columnar Mesophases Utilizing C2 Symmetric Aromatic Donor-Acceptor Complexes. *J. Am. Chem. Soc.* **2006,** *128* (24), 7995-8002.

85. Erukulla, R. K.; Byun, H. S.; Locke, D. C.; Bittman, R., Stereospecific and Regiospecific Ring-Opening of Glycidol with Primary and Secondary Alcohols Mediated by Diisobutylaluminium Hydride. *Journal of the Chemical Society-Perkin Transactions 1* **1995,** (18), 2199-2200.

86. Kitaori, K.; Furukawa, Y.; Yoshimoto, H.; Otera, J., CsF in organic synthesis. Regioselective nucleophilic reactions of phenols with oxiranes leading to enantiopure beta-blockers. *Tetrahedron* **1999,** *55* (50), 14381-14390.

87. Tambara, K.; Ponnuswamy, N.; Hennrich, G.; Pantos, G. D., Microwave-Assisted Synthesis of Naphthalenemonoimides and N-Desymmetrized Naphthalenediimides. *J. Org. Chem.* **2011,** *76* (9), 3338-3347.

88. Bolchi, C.; Pallavicini, M.; Fumagalli, L.; Moroni, B.; Rusconi, C.; Valoti, E., Univocal syntheses of 2-and 3-hydroxymethyl-2,3-dihydro[1,4]dioxino[2,3-b]pyridine enantiomers. *Tetrahedron-Asymmetry* **2005,** *16* (20), 3380-3384.

89. Bruckner, J. R.; Porada, J. H.; Dietrich, C. F.; Dierking, I.; Giesselmann, F., A Lyotropic Chiral Smectic C Liquid Crystal with Polar Electrooptic Switching. *Angewandte Chemie-International Edition* **2013,** *52* (34), 8934-8937.

90. Galli, C., Cesium Ion Effect and Macrocyclization - a Critical-Review. *Org. Prep. Proced. Int.* **1992,** *24* (3), 285-&.

91. Billman, J. H.; Parker, E. E., Amino acids. I. Glycine. *J. Am. Chem. Soc.* **1943,** *65*, 761-2.

92. Schoning, K. U.; Scholz, P.; Wu, X. L.; Guntha, S.; Delgado, G.; Krishnamurthy, R.; Eschenmoser, A., The alpha-L-threofuranosyl-(3 '-2 ')-oligonucleotide system ('TNA'): Synthesis and pairing properties. *Helv. Chim. Acta* **2002,** *85* (12), 4111-4153.

93. Ebert, M. O.; Mang, C.; Krishnamurthy, R.; Eschenmoser, A.; Jaun, B., The structure of a TNA-TNA complex in solution: NMR study of the octamer duplex derived from alpha-(L)-threofuranosyl-(3'-2')-CGAATTCG. *J. Am. Chem. Soc.* **2008,** *130* (45), 15105-15.

94. Horhota, A.; Zou, K.; Ichida, J. K.; Yu, B.; McLaughlin, L. W.; Szostak, J. W.; Chaput, J. C., Kinetic analysis of an efficient DNA-dependent TNA polymerase. *J. Am. Chem. Soc.* **2005,** *127* (20), 7427-34.

95. Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O., Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **1991,** *254* (5037), 1497-500.

96. Rasmussen, H.; Kastrup, J. S.; Nielsen, J. N.; Nielsen, J. M.; Nielsen, P. E., Crystal structure of a peptide nucleic acid (PNA) duplex at 1.7 A resolution. *Nat. Struct. Biol.* **1997,** *4* (2), 98-101.

97. Nielsen, P. E., Targeting double stranded DNA with peptide nucleic acid (PNA). *Current Medicinal Chemistry* **2001,** *8* (5), 545-550.

98. Bentin, T.; Larsen, H. J.; Nielsen, P. E., Combined triplex/duplex invasion of double-stranded DNA by "tail-clamp" peptide nucleic acid. *Biochemistry* **2003,** *42* (47), 13987-95.

99. Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A., Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* **1990,** *343* (6253), 33-7.

100. Switzer, C.; Moroney, S. E.; Benner, S. A., Enzymatic incorporation of a new base pair into DNA and RNA. *J. Am. Chem. Soc.* **1989,** *111* (21), 8322-3.

101. Liu, H.; Gao, J.; Lynch, S. R.; Saito, Y. D.; Maynard, L.; Kool, E. T., A four-base paired genetic helix with expanded size. *Science* **2003,** *302* (5646), 868-71.

102. Liu, H.; Lynch, S. R.; Kool, E. T., Solution structure of xDNA: a paired genetic helix with increased diameter. *J. Am. Chem. Soc.* **2004,** *126* (22), 6900-5.

103. Krueger, A. T.; Lu, H.; Lee, A. H.; Kool, E. T., Synthesis and properties of size-expanded DNAs: toward designed, functional genetic systems. *Acc. Chem. Res.* **2007,** *40* (2), 141-50.

104. Lee, A. H. F.; Kool, E. T., Exploring the limits of DNA size: Naphtho-homologated DNA bases and pairs. *J. Am. Chem. Soc.* **2006,** *128* (28), 9219-9230.

105. Matsuda, S.; Leconte, A. M.; Romesberg, F. E., Minor groove hydrogen bonds and the replication of unnatural base pairs. *J. Am. Chem. Soc.* **2007,** *129* (17), 5551-5557.

106. Yu, C. Z.; Henry, A. A.; Schultz, P. G.; Romesberg, F. E., Polymerase recognition of unnatural base pairs (vol 41, pg 3841, 2002). *Angewandte Chemie-International Edition* **2002,** *41* (24), 4612-4612.

107. Leconte, A. M.; Hwang, G. T.; Matsuda, S.; Capek, P.; Hari, Y.; Romesberg, F. E., Discovery, Characterization, and Optimization of an Unnatural Base Pair for Expansion of the Genetic Alphabet. *J. Am. Chem. Soc.* **2008,** *130* (7), 2336-2343.

108. O'Neill, B. M.; Ratto, J. E.; Good, K. L.; Tahmassebi, D. C.; Helquist, S. A.; Morales, J. C.; Kool, E. T., A highly effective nonpolar isostere of deoxyguanosine: synthesis, structure, stacking, and base pairing. *J. Org. Chem.* **2002,** *67* (17), 5869-75.

109. Guckian, K. M.; Krugh, T. R.; Kool, E. T., Solution structure of a DNA duplex containing a replicable difluorotoluene-adenine pair. *Nat. Struct. Biol.* **1998,** *5* (11), 954-9.

110. Guckian, K. M.; Schweitzer, B. A.; Ren, R. X.; Sheils, C. J.; Tahmassebi, D. C.; Kool, E. T., Factors Contributing to Aromatic Stacking in Water: Evaluation in the Context of DNA. *J. Am. Chem. Soc.* **2000,** *122* (10), 2213-2222.

111. Khakshoor, O.; Wheeler, S. E.; Houk, K. N.; Kool, E. T., Measurement and Theory of Hydrogen Bonding Contribution to Isosteric DNA Base Pairs. *J. Am. Chem. Soc.* **2012,** *134* (6), 3154-3163.

112. Moran, S.; Ren, R. X.; Rumney, S.; Kool, E. T., Difluorotoluene, a Nonpolar Isostere for Thymine, Codes Specifically and Efficiently for Adenine in DNA Replication. *J. Am. Chem. Soc.* **1997,** *119* (8), 2056-2057.

113. Brotschi, C.; Mathis, G.; Leumann, C. J., Bipyridyl- and biphenyl-DNA: a recognition motif based on inter-strand aromatic stacking. *Chem. - Eur. J.* **2005,** *11* (6), 1911-1923.

114. Zahn, A.; Leumann, C. J., Recognition properties of donor- and acceptor-modified biphenyl-DNA. *Chemistry* **2008,** *14* (4), 1087-94.

115. Grigorenko, N. A.; Leumann, C. J., Electron transfer through a stable phenanthrenyl pair in DNA. *Chem. Commun. (Cambridge, U. K.)* **2008,** (42), 5417-5419.

116. Baumstark, D.; Wagenknecht, H. A., Fluorescent hydrophobic zippers inside duplex DNA: interstrand stacking of perylene-3,4:9,10-tetracarboxylic acid bisimides as artificial DNA base dyes. *Chemistry* **2008,** *14* (22), 6640-5.

117. Wilson, J. N.; Gao, J. M.; Kool, E. T., Oligodeoxyfluorosides: strong sequence dependence of fluorescence emission. *Tetrahedron* **2007,** *63* (17), 3427-3433.

118. Gao, J.; Strassler, C.; Tahmassebi, D.; Kool, E. T., Libraries of composite polyfluors built from fluorescent deoxyribosides. *J. Am. Chem. Soc.* **2002,** *124* (39), 11590-1.

119. Dai, N.; Teo, Y. N.; Kool, E. T., DNA-polyfluorophore excimers as sensitive reporters for esterases and lipases. *Chem. Commun.* **2010,** *46* (8), 1221-1223.

120. Yuen, L. H.; Franzini, R. M.; Wang, S. L.; Crisalli, P.; Singh, V.; Jiang, W.; Kool, E. T., Pattern-Based Detection of Toxic Metals in Surface Water with DNA Polyfluorophores**. *Angewandte Chemie-International Edition* **2014,** *53* (21), 5361-5365.

121. Kwon, H.; Jiang, W.; Kool, E. T., Pattern-based detection of anion pollutants in water with DNA polyfluorophores. *Chemical Science* **2015,** *6* (4), 2575-2583.

122. Zheng, Y.; Long, H.; Schatz, G. C.; Lewis, F. D., Duplex and hairpin dimer structures for perylene diimide-oligonucleotide conjugates. *Chem. Commun. (Cambridge, U. K.)* **2005,** (38), 4795-4797.

123. Wang, W.; Wan, W.; Zhou, H.-H.; Niu, S.; Li, A. D. Q., Alternating DNA and π-conjugated sequences. Thermophilic foldable polymers. *J. Am. Chem. Soc.* **2003,** *125* (18), 5248-5249.

124. Endo, M.; Fujitsuka, M.; Majima, T., Diastereochemically Controlled Porphyrin Dimer Formation on a DNA Duplex Scaffold. *J. Org. Chem.* **2008,** *73* (3), 1106-1112.

125. Endo, M.; Shiroyama, T.; Fujitsuka, M.; Majima, T., Four-way-branched DNA-porphyrin conjugates for construction of four double-helix-DNA assembled structures. *J. Org. Chem.* **2005,** *70* (19), 7468-7472.

126. Yen, S. F.; Gabbay, E. J.; Wilson, W. D., Interaction of aromatic imides with deoxyribonucleic acid. Spectrophotometric and viscometric studies. *Biochemistry* **1982,** *21* (9), 2070-6.

127. Cuenca, F.; Greciano, O.; Gunaratnam, M.; Haider, S.; Munnur, D.; Nanjunda, R.; Wilson, W. D.; Neidle, S., Tri- and tetra-substituted naphthalene diimides as potent G-quadruplex ligands. *Bioorg. Med. Chem. Lett.* **2008,** *18* (5), 1668-1673.

128. Di Antonio, M.; Doria, F.; Richter, S. N.; Bertipaglia, C.; Mella, M.; Sissi, C.; Palumbo, M.; Freccero, M., Quinone Methides Tethered to Naphthalene Diimides as Selective G-Quadruplex Alkylating Agents. *J. Am. Chem. Soc.* **2009,** *131* (36), 13132-13141.

129. Fathi, R.; Rudolph, M. J.; Gentles, R. G.; Patel, R.; MacMillan, E. W.; Reitman, M. S.; Pelham, D.; Cook, A. F., Synthesis and Properties of Combinatorial Libraries of Phosphoramidates. *J. Org. Chem.* **1996,** *61* (16), 5600-5609.

130. Tinoco, I., Hypochromism in Polynucleotides. *J. Am. Chem. Soc.* **1960,** *82* (18), 4785-4790.

131. Friedman, R. A.; Honig, B., A Free-Energy Analysis of Nucleic-Acid Base Stacking in Aqueous-Solution. *Biophys. J.* **1995,** *69* (4), 1528-1535.

132. Honig, B.; Sharp, K.; Yang, A. S., Macroscopic Models of Aqueous-Solutions - Biological and Chemical Applications. *J. Phys. Chem.* **1993,** *97* (6), 1101-1109.

133. Saenger, W., *Principles of nucleic acid structure*. Springer-Verlag: New York, 1984; p xx, 556 p.

134. Senior, M.; Jones, R. A.; Breslauer, K. J., Influence of Dangling Thymidine Residues on the Stability and Structure of 2 DNA Duplexes. *Biochemistry* **1988,** *27* (10), 3879-3885.

135. Crothers, D. M.; Zimm, B. H., Theory of the Melting Transition of Synthetic Polynucleotides: Evaluation of the Stacking Free Energy. *J. Mol. Biol.* **1964,** *9*, 1-9.

136. Devoe, H.; Tinoco, I., Jr., The stability of helical polynucleotides: base contributions. *J. Mol. Biol.* **1962,** *4*, 500-17.

137. Gray, D. M.; Tinoco, I., A New Approach to Study of Sequence-Dependent Properties of Polynucleotides. *Biopolymers* **1970,** *9* (2), 223-&.

138. SantaLucia, J., A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences of the United States of America* **1998,** *95* (4), 1460-1465.

139. Allawi, H. T.; SantaLucia, J., Jr., Thermodynamics and NMR of internal G.T mismatches in DNA. *Biochemistry* **1997,** *36* (34), 10581-94.

140. Gray, D. M., Derivation of nearest-neighbor properties from data on nucleic acid oligomers. II. Thermodynamic parameters of DNA.RNA hybrids and DNA duplexes. *Biopolymers* **1997,** *42* (7), 795-810.

141. Petruska, J.; Goodman, M. F., Enthalpy-entropy compensation in DNA melting thermodynamics. *J. Biol. Chem.* **1995,** *270* (2), 746-50.

142. Chaires, J. B., Possible origin of differences between van't Hoff and calorimetric enthalpy estimates. *Biophys. Chem.* **1997,** *64* (1-3), 15-23.

143. Liu, Y.; Sturtevant, J. M., Significant discrepancies between van't Hoff and calorimetric enthalpies. III. *Biophys. Chem.* **1997,** *64* (1-3), 121-6.

144. He, Y. C.; Shang, Y. Z.; Liu, Y.; Zhao, S. L.; Liu, H. L., Melting dynamics of short dsDNA chains in saline solutions. *Springerplus* **2015,** *4*.

145. Wong, K. Y.; Pettitt, B. M., The Pathway of Oligomeric DNA Melting Investigated by Molecular Dynamics Simulations. *Biophys. J.* **2008,** *95* (12), 5618-5626.

Vita

Brian Aki Ikkanda was born and raised in sunny southern California. It was in fourth grade where his passion for science began and by fifth grade he developed an interest exploring the world of chemistry. After attending Ayala High School, he attended the University of California, Berkeley where he would continue his studies in chemistry. He had the opportunity to do research under the guidance of Dr. Richmond Sarpong researching the total synthesis of natural products and also Dr. Isao Kubo researching non-natural amino acids extracted from plants. He then had the opportunity to do research under the guidance of Dr. Ron Zuckermann at Lawrence Berkeley National Laboratories investigating peptoid nanosheets. In the fall of 2009, he began his graduate studies under Dr. Brent Iverson where his passion for teaching and chemistry was further cultivated. One of his favorite life verses is Proverbs 3:5-6.

[5] Trust in the Lord with all your heart,
and do not lean on your own understanding.
[6] In all your ways acknowledge him,
and he will make straight your paths.
~ Proverbs 3:5-6

Email Address: bikkanda@utexas.edu

This dissertation was typed by Brian Aki Ikkanda.